UNIVERSITY OF TRENTO
Physics Department

Production and excitation of cold Ps for $\bar{H}$ formation by charge exchange: towards a gravitational measurement on antimatter

PhD Thesis

Francesco Guatieri

UNIVERSITY OF TRENTO

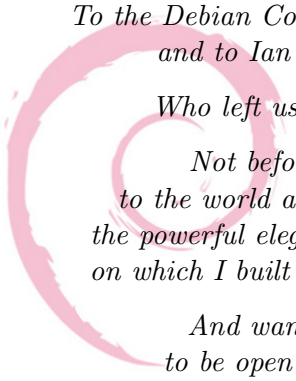
PhD School in Physics
30^{th} Cycle

Production and excitation of cold Ps for $\bar{H}$ formation by charge exchange: towards a gravitational measurement on antimatter

Tutor:
Prof. Roberto Sennen Brusa

Referees:
Prof. Christoph Hugenschmidt
Dr. László Liszkay

Candidate:
Francesco Guatieri



To the Debian Community
and to Ian Murdock

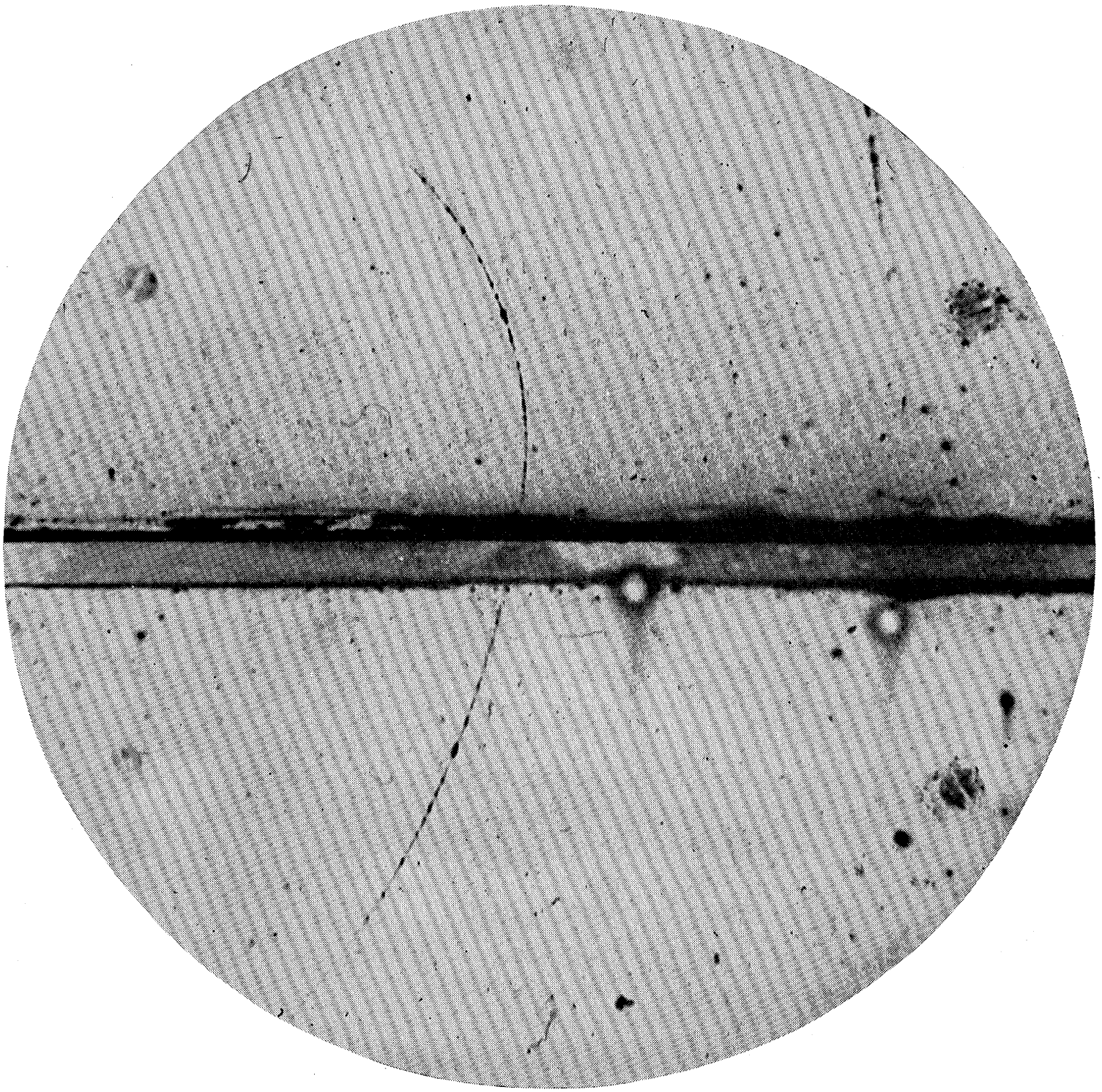
Who left us in 2015,

Not before gifting
to the world and to me
the powerful elegant tools
on which I built my work

And wanted them
to be open and free,
rarest of gems
nowadays

Contents

On the back of this page, the track that allowed the discovery of the positron, published by Carl Anderson in his 1932 work[1]



Introduction

– Of these four forces, there's one we don't really understand.
– Is it the weak force or the strong–
– It's gravity.

– Randall Munroe

The 2nd of August 1932 in a sun-blessed[2] Pasadena, Carl D. Anderson is working at his cloud chamber located on the top floor of the aeronautics building at Caltech[3] when he shoots the historic photograph (shown in the facing page) that marks the discovery of the positron[1]. This accomplishment, which will earn Anderson the Nobel prize and initiate the whole field of antimatter investigation, marks the first time in history in which antimatter is ever observed. Consequently it is also the first time antimatter is seen interacting through one of the four fundamental forces; in the case of Anderson's chamber the electromagnetic.

Shortly after his first historic discovery another milestone will be set when in 1936 Anderson and Neddermeyer observe particles in the cosmic rays that appear to have a mass comprised between that of an electron and that of a proton[4]: it is the first detection of muons and antimuons. In modern terms we know that the dominant decay of charged muons is mediated by $W^{\pm}$ bosons, thus making the detection of the spontaneous muon decay in 1940[5] the first instance in which antimatter is observed to interact through the weak force.

In 1949 De Benedetti discovers that the emission angle of the gamma rays originating from positron annihilations do not form an exact 180 degree angle[6], laying the foundation for the investigation of the electronic structure of matter based on ACAR (Angular Correlation of Annihilation Radiation). In the early days of positronics the mainstream paradigm was to employ the interaction of positrons with matter to study the properties of positrons; in

the early '50s the use of positrons to study the properties of matter begins its traction.

In 1950 R. Bjorklund, W. E. Crandall, B. J. Moyer and H. F. York while bombarding solid targets with high energy protons detect the production of hard gamma rays that could be interpreted as the decay products of a novel extremely short-lived meson[7]; this is the first experimental evidence of the existence of the neutral pion whose production by proton bombardment of nuclei is mediated by the strong interaction. At this point antimatter has been seen to interact with three of the four fundamental forces (gravity being the missing one).

The next year Martin Deutsch observes a bound state of an electron and positron[8] and successively with Everett Dulit determines its lifetime to depend on the relative spin orientation of its two constituent particles[9]. Positronium (Ps) has been discovered. Meanwhile medical science, whose path has already crossed that of particle physics with medical employment of X-rays, is getting interested in possible applications of the newly-discovered antimatter. In 1953 Brownell and Sweet[10] demonstrate the potential of positron annihilation to provide imaging of brain tumors and the first prototype of the PET (Positron Emission Tomography) scanner is built.

In 1955 the antiproton ($\bar{\mathrm{p}}$) is discovered by Emilio Segrè and Owen Chamberlain who produces it by bombarding a copper target with a 4.3 GeV proton beam[11].

Next year Wu's famed experiment[12] shows violation of the parity symmetry by the weak interaction, leading to the assumption of the fundamental symmetry underlying fundamental interactions to be CPT (Charge, Parity and Time-reversal). At the same time works by Pauli and Luders[13] determines the CPT symmetry to be predicted by a wide variety of Quantum Field Theory (QFT) models at the tree level (more general proofs of the CPT theorem will come in the early '60s). Consequence of the CPT symmetry is that mass and mean life of particles and their antiparticles should be equal, allowing as a consequence the formulation of high precision experimental tests over the interaction of antimatter through the fundamental forces.

In 1960 Bell and Jørgensen discover that the lifetime of positrons implanted in solids depends on the material characteristics[14]. A new field of material investigation and defect studies through the assessment of the positron lifetime takes off. The next year the first PET scanner designed for clinical purpose is built[15]; ten years later, in 1971 the first computerized PET apparatus is born [15]. The increased resolution and ease of use, alongside with the development of medically-suitable radionuclides markers, will lead to the spread of PET scan adoption in the '70s.

In 1978 through meticulous preparation of clean metal surfaces Mills produces the first moderated ($< 5\,\mathrm{eV}$) positron beam[16]. Three years later Schwinberg, Van Dyck and Dehmelt are able to capture $1.3 \cdot 10^7$ positrons

inside of a Penning trap[17]. In 1983, LEAR (Low Energy Antiproton Ring) is commissioned, providing the first slow ($< 50\,\mathrm{MeV}$) antiproton beam in history[18]. A whole new range of possibilities to experiment with cold antimatter becomes available.

In 1984 the record for the CPT test through measurement of the mass difference of $e^{\pm}$ is set by Chu, Mills and Hall[19] who determine $|(m_{e^+} - m_{e^-})/m_{e^-}| < 40 \cdot 10^{-9}$ (parts per billion). In the same year the record CPT test by $\mu^{\pm}$ lifetime difference is set by Bardin et al.[20] who determine it to be less than $10 \cdot 10^{-6}$.

In 1989 Surko improves the design of the positron trap[21] by introducing a carefully-controlled amount of residual gas in the chamber and allowing the trapped positrons to be cooled by their interaction with it. Surko's design will become the workhorse of many modern slow positron experiments.

In 1990 the absolute record test over the CPT symmetry is set at CERN[22] when the mass difference between the neutral kaon and its antiparticle is determined to be $|(m_{K^0} - m_{\overline{K^0}})/m_{K^0}| < 5 \cdot 10^{-18}$.

In 1995, employing the slow antiproton beam provided by LEAR to bombard a Xe target, the PS210 successfully produces, for the first time in history, antihydrogen ($\bar{\mathrm{H}}$)[23].

In 2000 the Antiproton Decelerator (AD), successor of LEAR, begins operation; its commissioning initiates a new era of experiments involving cold baryonic antimatter. To give some examples, employing its beam in 2002 the ATHENA collaboration was able to produce abundant (> 50000) quantities of $\bar{\mathrm{H}}$ [24]. In 2011 the ALPHA collaboration is able to trap anti-hydrogen produced from the AD beam for more than 1000 s [25]. In the same year Hori et al. determine precisely the mass ratio between the electron and antiproton[26], the result translates into test of CPT invariance in the form of a limit to the relative mass difference between proton and antiproton of $|(m_p - m_{\bar{p}})/m_p| < 0.7 \cdot 10^{-9}$.

Today. After 85 years of scrupulous investigation, high precision measurements and technological employment of antimatter no one has yet been able to observe it interact through the gravitational interaction.

Our most solid theory of gravitation, general relativity, postulates the universality of the weak equivalence principle, consequence of which is that antimatter and matter should be identically subject to the gravitational force. The best experimental upper bound available today on the ratio between gravitational mass m_g and inertial mass m_i for antimatter has been given by the ALPHA experiment in 2013 as $|\,m_g/m_i| < 75$ [27] while to our best knowledge no lower bound exist on the ratio $|\,m_g/m_i|$. For comparison the best tests of the weak equivalence principle for ordinary matter were able to determine the aforementioned ratio to be constant up to more that 0.21 parts per trillion [28].

The aim of the AE$\bar{\mathrm{g}}$IS experiment is to measure directly the gravitational

pull on an uncharged probe, antihydrogen, to determine the ratio between its gravitational and inertial mass with an uncertainty inferior to 10^{-2}[29].

The AEgIS experiment

Born in 2008 out of a collaboration that currently comprises 18 institutes from all around Europe, AEgIS is one of the six experiments located inside of the AD hall at CERN[29].

The aim of the AEgIS experiment is to produce large quantities of cold antihydrogen to then accelerate it towards a Moirè deflectometer[30] located in a highly electromagnetically shielded region. The geometry of the deflectometer (see as a reference figure 1) allows us to compute the gravitational pull that the $\bar{H}$ atoms are subjected to as:

$$g = 4 \cdot \frac{\Delta y}{\Delta t^2} \tag{1}$$

where Δt is the time required by the $\bar{H}$ to traverse the entire deflectometer and Δy is the vertical displacement it was subjected to in flight[31]. Measuring gravitational pull through the deflectometer does not require the employment of a collimated or monochromatic beam; nonetheless it raises other two fundamental requirements. The first is on the pulsed nature of the beam, which is required to measure the time of flight (ToF) giving the measurement of Δt. The second requirement concerns the temperature of the employed $\bar{H}$. The vertical deflection (as patently stated in equation 1) depends primarily on the ToF, which is limited by the antihydrogen temperature. To our best estimates the detection of 1000 $\bar{H}$ with a temperature of 0.1 K would permit the achievement of the goal of AEgIS to measure gravitational pull with 1% uncertainty[32].

Producing cold, abundant and pulsed antihydrogen makes many of the techniques employed previously to produce $\bar{H}$ unamenable. $\bar{H}$ production by two-γ conversion (as done by Baur et al.[23]) is inefficient and would yield extremely energetic antihydrogen. The direct mixing method employed by ATHENA[24], which consists in loading a trap with e^+ and then in having an antiproton bunch cross the positron plasma, has better efficiency and produces colder $\bar{H}$, still it doesn't meet the requirements of the AEgIS experiment. The main limitation of mixing is due to the difference in charge between $\bar{p}$ and e^+ which creates a deep potential well that accelerates particles prior to collision, originating energetic reaction products and leading to high re-ionization rates.

A way to circumvent this limitation is to use, as AEgIS aims to, the charge exchange reaction[33, 34] depicted in figure 2. Charge exchange requires a plasma of antiprotons to be crossed by a cloud of positronium (Ps) excited

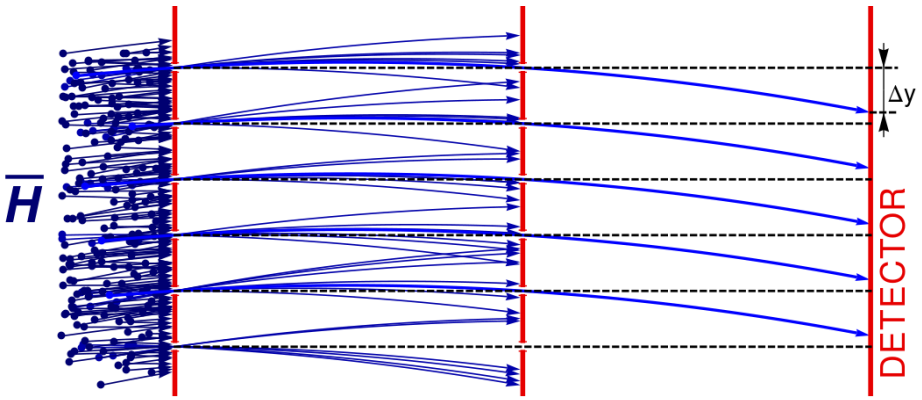


Figure 1: Working principle of a Moiré deflectometer: the first reticle selects particles based on their position, the second one selects them according to their initial velocity. This is indeed an oversimplification of what the actual design of the AE$\bar{g}$IS detector will be. Being the detector the last component that needs to be assembled onto the apparatus, a specialized team in the AE$\bar{g}$IS collaboration spent the last years to develop much more sophisticated designs capable of employing diffraction to increase precision. The final design will be, indeed, very similar to a Talbot-Lau interferometer.

into high Rydberg states. Employment of charge exchange grants a low re-ionization rate, the possibility to control which quantum state $\bar{H}$ is produced in (by controlling the impinging Ps quantum state) and the fact that the final temperature is dominated by the $\bar{p}$ temperature. By employing cold antiprotons and a pulsed positronium source it should be, in principle, possible to meet the demands of the gravitational measurement.

The main body of AE$\bar{g}$IS (see figure 3) consists in two cylindrical superconducting magnets which generate a field of 1 T and 4.46 T respectively and that are traversed by a beam lying on their axis. The beam line is riddled with cylindrical electrodes composing a total of 6 Malmberg-Penning traps (see the zoomed portion of figure 3). The end of the beam line exiting the 4.46 T magnet connects to the AD apparatus to receive antiprotons. The other end will connect to the antigravity detection apparatus (more on that later). Upstream of the point where the beam line enters the 4.46 T magnet, another beam line merges onto the primary one: it is the line dedicated to the production of a pulsed positron beam. Sitting on the side of the experiment, a laser hut hosts three laser systems whose beams can be routed in the region that hosts the final trap of the system inside of the 1 T magnet.

The operation of AE$\bar{g}$IS starts when a bunch of $\bar{p}$ is received from AD. Antiprotons from AD are still highly energetic (≈ 5 MeV); upon being received by AE$\bar{g}$IS they are moderated by having them traverse a silicon-

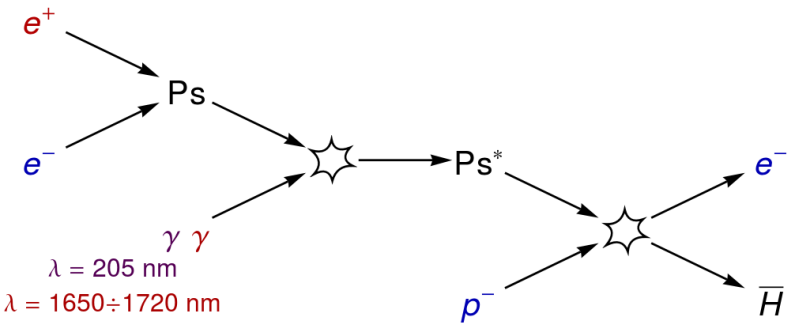


Figure 2: Production of $\bar{\mathrm{H}}$ by charge exchange. A bound state of positron and electron (positronium) is excited into Rydberg levels and let interact with a cloud of antiprotons. During the interaction the positron is lost by the positronium and acquired by the antiproton that, in turn, becomes an antihydrogen.

aluminum degrader which reduces their energies to a continuum spectrum mostly comprised below 100 keV, of these the fraction whose energy is below 9 keV is then captured inside a combination of the first two traps in the AEgIS apparatus which are named C and P and located inside of the 4.46 T magnet. The moderation and capture process is performed with an overall efficiency of about $1 \div 2\%$.

Antiprotons stored in the C+P trap are then cooled down by their interaction with a previously loaded electron cloud and, as their energy decreases, they are moved inside of the P trap which features rotating wall electrodes. The antiproton bunch is then further compressed by means of the rotating walls[35] in preparation of their transfer into the final trap of the apparatus: the *Production* trap, located inside of the 1 T magnet. The additional compression must account for the production trap's small radius and for the increase in the plasma radius caused by its transfer into a weaker magnetic field. The process might be repeated several times to stack several AD shots and increase the number of $\bar{\mathrm{p}}$ accumulated in the production trap.

While antiprotons are held in the production trap a shot of positrons is transferred from the positron line across the entire apparatus and onto a positron-positronium converter (more on that later). A cloud of cold positronium is emitted by the converter and enters the production trap through segmented electrodes to cross the $\bar{\mathrm{p}}$ plasma. As positronium is flying towards the production trap, excitation lasers are shone onto the cloud exciting Ps atoms in Rydberg states up to $n = 18$. As Rydberg positronium meets cold antiprotons, excited $\bar{\mathrm{H}}$ is produced. $\bar{\mathrm{H}}$ will then be accelerated by means of Stark acceleration[36, 37] towards the Moiré deflectometer to perform the gravitational measurement.

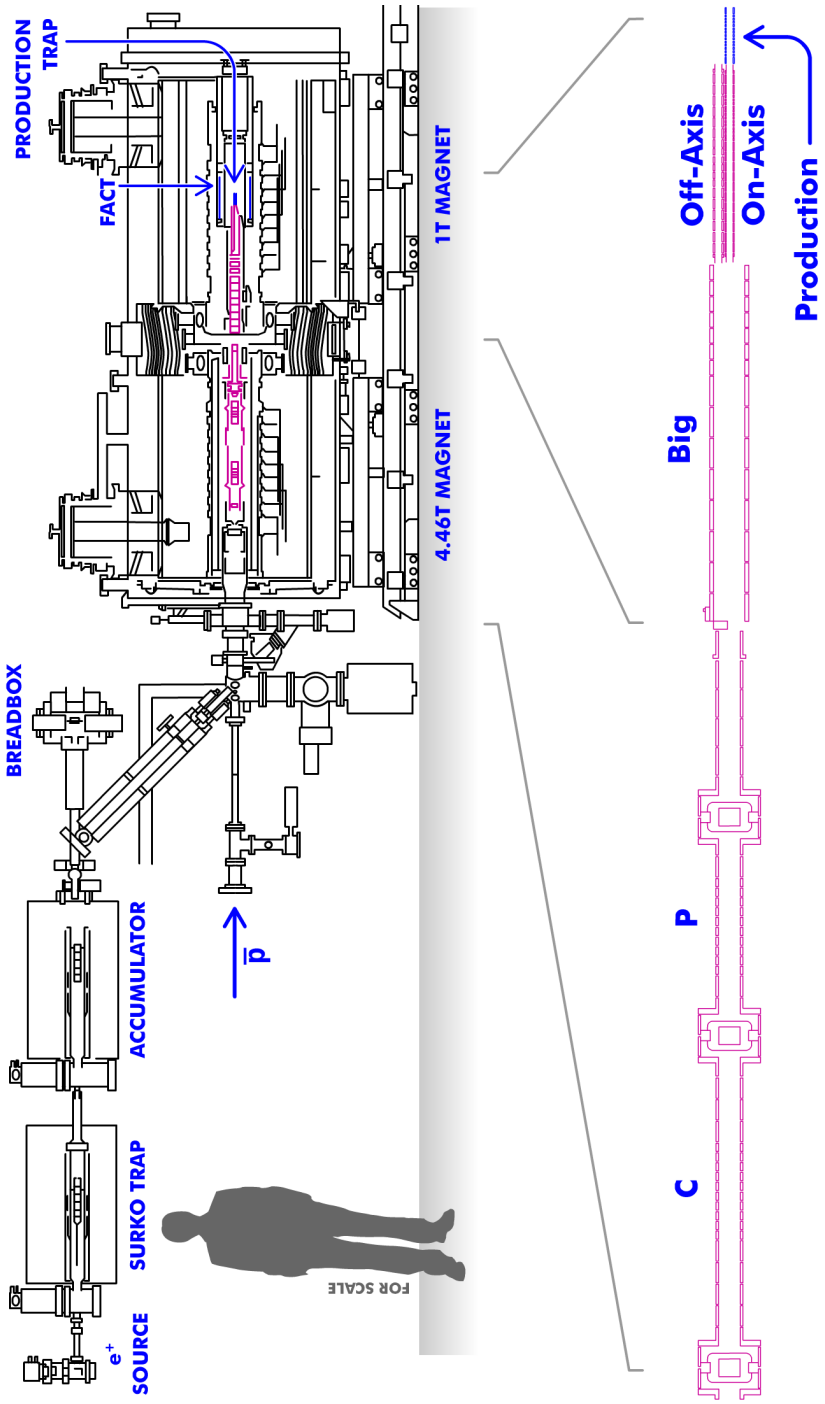


Figure 3: On the top (relatively to the sideways direction) a schematic depiction of the AE$\bar{g}$IS apparatus, below a zoom in of its main trap system, the labels indicating each trap denomination.

Positrons

We mentioned the presence of a positron line and, generically, its capability to produce shots of positrons. In chapter 2 we'll discuss the use of another positron line: SURF[38] located in the physics department of the University of Trento. We'll introduce here some general facts regarding positron lines in general and in the specific case of the aforementioned two.

As of today we can divide the technologically feasible positron sources in three categories. The first is pair production in which a target is bombarded with high-energy electrons and positrons are produced in the resulting shower[39]. The second is to produce positrons by pair production from γ rays, which in turn are produced from neutron conversion, this is the method employed e.g. at the NEPOMUC[40] facility. Finally it is possible to employ β^+-decaying radioactive sources such as ^{22}Na which are produced off-site and need to be replaced every few years. The first two methods require large infrastructures (namely linacs and nuclear reactors) whose size, cost and logistic difficulty made them unfeasible to be employed either inside the AE$\bar{\text{g}}$IS experimental zone or in the positron laboratory of the University of Trento. Therefore the design of both lines needed to rely on the third possibility.

At the beginning of both positron lines sits a ^{22}Na source with an activity of $\approx 50\,\text{mCi}$. Positrons coming from β^+ decays exhibits a continuous spectrum whose endpoint energy, in the case of ^{22}Na, is 546 keV[41]. Before they can be trapped the energy of positrons must be lowered to a few eV at most. This operation is called moderation and was pioneered by Mills[16]. Moderation of positrons can be performed by letting them impinge on a thin metal sheet (as is done in SURF using a 1 µm thick tungsten foil) which moderates them in the transmission scheme. Instead if the maximum efficiency is desired and a cryogenerator is available (which is the case of AE$\bar{\text{g}}$IS) the moderator of choice is solid neon (once again pioneered by Mills et al.[42]); in this case the moderation takes place in the reflection scheme. In both cases the main challenge of positron moderation is maintaining the moderator material devoid of impurities which could capture positrons impeding their re-emission in vacuum. In the case of metal moderators, this is usually accomplished by baking them to the highest possible temperature in ultra high vacuum, in the case of neon moderator this is achieved by evaporating and growing back periodically the moderator.

Here the architecture of the two beams diverges widely since the SURF beam is continuous while that produced in the AE$\bar{\text{g}}$IS positron line is pulsed.

To produce a continuous beam, SURF features an electrostatic deflector that bends the beam by 90 degrees, getting rid, thus, of the non moderated positrons. Then a focusing stage follows culminating in a final acceleration stage that gives the beam its final energy. The pulsed beam in AE$\bar{\text{g}}$IS is obtained, instead, with two consecutive trapping and accumulation stages

which accumulate and further cool positrons from the source for a time that can span up to several minutes depending on the required beam intensity. When the accumulation phase is concluded the positrons are then ejected all at once from the accumulator trap and directed into the beam line to be used in the experiment.

Continuous positron beams can be obtained from more compact apparatuses and require less expensive equipment than pulsed ones. Two main advantages can be recognized, conversely, for the employment of a pulsed beam. The first is the generation, as a direct result of its operation, of a start time which is valuable in annihilation spectrum or time of flight (TOF) measurements. The second is that when accumulation and cooling stages are integrated in a pulsed positron line (as is the case of AE$\bar{g}$IS) much higher intensities can be obtained, which allows pulsed lasers to be synchronized with the operation of the positron apparatus, resulting potentially in extremely high excitation efficiencies[1]. While a continuous beam is usually well-suited for material studies, a pulsed one is required for the operation of AE$\bar{g}$IS.

Positronium

Positronium is a bound state of an electron and a positron. It is intrinsically short lived due to the absence of any conservation law preventing its components from annihilating into two or three γ rays. Depending on the spin configuration of its constituents this can characteristically happen in 125 ps (in the case of para-positronium) or in 142 ns (for ortho-positronium).

On top of its role in the charge exchange reaction employed in AE$\bar{g}$IS, positronium might provide high accuracy QED tests. Positronium being a pure leptonic system does not require the QCD corrections necessary to compute with high precision the energy level of the hydrogen atom[2]; our limited knowledge of which is closing in as the limitation to experimental verification of QED through spectroscopy[43].

To the first order of approximation the energy levels of positronium are the same as those of the hydrogen atom, with their energy halved due to the reduced mass of the e^+e^- system being approximately half that of the pe^- system. Excited levels of the positronium atom tend to be longer lived, with the lifetime of high Rydberg states reaching up to tens of µs[44].

Upon the implantation of positrons, positronium is spontaneously formed at the surface of several solid materials as it is reached by a thermalized positron[45] and in the bulk of some insulators while positrons diffuse in

[1] Moreover extremely high beam intensities open the possibility for the performance of a whole new spectrum of physics experiments involving $e^+ - e^+$or Ps − Ps interactions

[2] Playing on the nuance of the term, a physicist could paradoxically argue that positronium is a true "hydrogen atom" while, due to QCD, actual hydrogen is not

them[45]. Both the production efficiency and the emission energy of positronium from a surface are strongly dependent on the material's work function. Ideally, whether to perform spectroscopy measurements or to produce antihydrogen, we would like to be able to produce an abundance of cold positronium; unluckily the materials that emit positronium at the lowest energies tend to be the least efficient converters and vice versa[46, 47]. The current technological solution to this impasse is to employ a material with a high Ps yield to produce positronium and then cool the Ps atoms afterwards.

A widely used positron to positronium converter is SiO_2. The history of its employment as a converter dates back to 1978 when Mills documented that silica-coated surfaces presented an excellent positronium yield[48], reaching over 40% for low implantation energies. The late '90s and the early 2000s saw a surge in the interest to study the behavior of positrons and positronium in porous silica with studies that analyzed positronium yield, lifetime of positronium in cavities and the 2γ to 3γ annihilation ratio (to give an example [49, 50, 51]). These studies were motivated on one side by the capability of the spectroscopy to investigate the characteristic of the pores in the material, on the other by the prospect of porous silica to provide a suitable environment in which a Ps Bose-Einstein condensate (BEC) could in the future be produced.

In 2003 Yu et al. detected positronium being emitted from porous silsesquioxane films with energies as low as 0.1 eV [52]. In 2008 Liszkay et al. showed that mesostructured silica bombarded with positrons was capable of emitting positronium even when the implantation energy was bound to implant the e^+ up to 300 nm inside the material[53]. Liszkay further observed that the energy of the emitted Ps presented a dependence on the implantation depth and suggested mesostructured silica as a candidate for the production of cold positronium to be used in the synthesis of $\bar{H}$[3]. In 2010 Mariazzi et al. demonstrated the capability of nanochanneled silicon converters to produce thermalized positronium[54].

Nanochanneled silicon converters (or NCP, nanochanneled plates) are silicon monocrystal chips on the surface of which nanoscopic channels (typically 2 µm long and $\approx$ 10 nm wide) have been electrochemically etched and their surface oxidized (two SEM images of NCPs are shown in figure 4). When positrons are implanted into the NCP they gradually lose their kinetic energy until they reach thermalisation. If energies are kept below 12 keV most of them will reach thermalisation in the region of the NCP traversed by the nanoscopic channels. Positrons then diffuse into the material until they reach the surface of a channel or annihilate in the bulk; due to the long diffusion distance of positrons in materials ($\approx$ 200 nm) and the high channel density of the sample (which can exceed $10^{16}/m^2$) a large portion of the implanted

[3] A more complete review on positron cooling inside of porous materials can be found in [56]

positrons reaches the surface of a nanoscopic channel before annihilation. A portion of the positrons that does so is then emitted inside of the channel itself as positronium with an energy of 3 eV. The positronium then travels along the channel until it either reaches the surface of the NCP or it annihilates. During its travel along the channel, positronium interacts with the channel walls losing energy to the material to be emitted, in the end, at a much lower energy than that of its production. Positronium with temperatures close to complete thermalisation has been observed to be emitted by NCPs[54, 55].

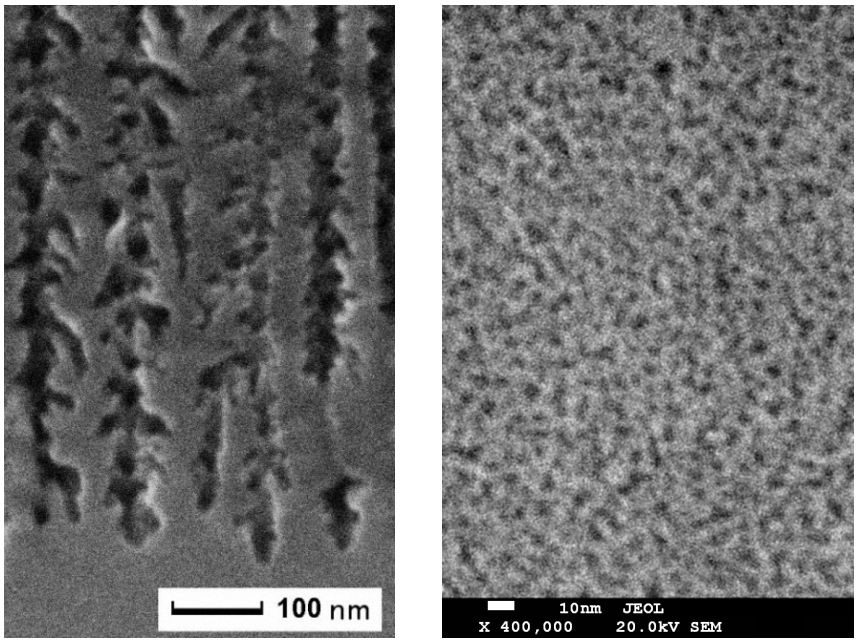


Figure 4: SEM pictures of an NCP. On the left a detail of a section view, on the right a top view. The section shown in the right panel displays secondary channels sprouting out from the primary ones, this happens when the etching process is tuned to produce larger channels and is not expected to take place in the samples that we are currently using and that will be tested in chapter 2; at the same time imaging thinner channels is difficult due to their structure collapsing when the sample is split to be SEM imaged.

Content of this thesis

To our knowledge no complete model of the physics of positronium production and thermalisation in nanochanneled plates has yet been proposed although being able to make prediction over the performance of silicon converters could potentially provide an invaluable guide in developing more robust and efficient

ones or efficient transmission targets. It has been speculated indeed several times[56, 57] that a completely classical model could fail to accurately describe the behavior of cold positronium inside of the nanoscopic channels; nonetheless this conjecture has never been verified. We'll spend chapter 1 to formulate a classical model of positronium production and thermalisation in NCPs and will validate it by testing it against the available experimental data.

In chapter 2 we will describe the measurement of the energy spectrum of positronium produced by nanochanneled plates using the beam produced by the SURF machine. Knowledge of the behavior of positronium leaving the nanochanneled plate is necessary in order to design precision physics experiments making use of NCP converters to produce positronium and for the optimization of the geometry of the $\bar{H}$ production region in AE$\bar{g}$IS. At the end of the chapter we compare the measured energy spectra data with the model proposed in chapter 1 showing, in the comparison, the indication of a transition during thermalisation process to a regime where quantum phenomena become significant.

On top of being a necessary tool for the production of $\bar{H}$, the AE$\bar{g}$IS positron line provides a beam that constitutes a valuable research resource by itself. Since the positron line commissioning, fundamental research of positrons and positronium physics has become an integral part of the AE$\bar{g}$IS scientific endeavor. During the course of the last three years, by employing the positron beam line of the experiment AE$\bar{g}$IS, we performed several positronium spectroscopy measurements that we'll describe in detail in chapter 3. Near the end of the chapter we will illustrate an improved version of the detrending technique commonly employed in signal analysis; our version, applied to the analysis of SSPALS, improves greatly the achievable precision of the experimental results.

The main goal of the production of positronium in AE$\bar{g}$IS is allowing for the production of antihydrogen, the detection of which poses some technical challenges. In chapter 4 we describe an innovative approach that we are currently pursuing to employ the detector FACT, part of the AE$\bar{g}$IS apparatus, to confirm the successful production of $\bar{H}$.

The four appendices offer some further reading into some of the details of the material exposed in the chapters. While the main chapters' focus is to discuss the physics of the proposed experiments and models, the appendixes will licentiously indulge in the technological details of the undertaking.

We'll include two final remarks before proceeding to the first chapter.

Firstly we avoided the use of the measurement unit Hz for Poisson-distributed signals. The reason is to increase clarity by expressing as much information as possible on the measured quantity through the choice of the measurement unit[4]. Although the rate of a Poisson-distributed signal, like the

[4]Consider, as an extreme case, how confusing if would be is a measurement of torque

frequency of a periodic signal, is the inverse of a time, most of the spectral analysis techniques that could be applied to periodic signals either lead to false conclusions or need special precautions when applied to Poisson-distributed ones. The International System of Units contains indeed a measurement unit specifically designed to describe Poisson-distributed signals: the Becquerel (Bq) which, unfortunately, has been officially defined only as suitable to describe the activity of a radionuclide. We'll take the licence to employ Bq more freely to indicate the rate of any Poisson-distributed signal.

Secondly modern experimental physics is, with few exceptions, a team work and, therefore, most experiments require teamwork to be carried out. My ideal goal, while writing this thesis, was to mention in it only work that was carried out by me personally. Unfortunately this choice would have required to skip some fundamental parts and would have severely hampered the comprehension of the text. This forced opting for the less radical solution of limiting the extent of these parts as much as possible. If some parts of the exposition feel rushed compared to the level of detail of the surrounding text, this is the reason.

were to be presented in Joules

Artistic rendering of a NCP geometry generated with the same procedures described in the chapter. To make the channel structure recognizable, surfaces have been rendered as translucent and surface color has been made dependent on the $\hat{y}$ coordinate. At the same time generation parameters have been altered so that channels are shorter, sparser and more tortuous. The image has been generated with the same code employed to perform Monte Carlo simulations.

Chapter 1

Simulating the Nanochanneled Plate

"No game," she said "Only a story for you to tell, so close to reality that it's almost indistinguishable from it."

– John McCrae
Twig

As detailed in the introduction, production of abundant cold positronium is paramount to the create cold $\bar{H}$ by charge exchange. Nanochanneled silica-coated silicon chips (hereafter nanochanneled plates or NCPs) have proved extremely effective converters of positron beams having energies of a few keV into positronium with energies of a fraction of an eV [54, 55]. Goal of the work presented in this chapter is to identify a suitable model to describe such conversion process; then to develop a simulation technique capable of reproduce the expected Ps velocity distribution as a function of the model parameters.

1.1 The Monte Carlo technique

A convenient way to describe the behavior of particles within macroscopic objects, as is in our case the slowing down of positrons inside of the NCP, is to model it as a series of stochastic discrete interactions that the impinging particles undergo while traversing the material. The Ps thermalisation process too, as long as it is treated as classical, which is our goal in this context, is effectively modeled by a series of discrete interactions taking place between the channel wall and the positronium atom.

The best way to compute the effects on a particle population undergoing a series of stochastic interactions is that of Monte Carlo simulation. In our specific application the Monte Carlo technique consists in sampling the impinging particles using an opportune distribution, then to simulate all of the relevant interactions we expect such particle to undergo, again drawing all of the interaction's random parameters from their respective distributions. We then record all of the parameters of the particle's final state we deem meaningful. The process is repeated on as many randomly drawn particle as possible and the resulting statistical distribution of the recorded parameters constitutes the result of the simulation process.

It is, in our case, convenient to divide the simulation process in two consecutive phases. The simulated particles will undergo a first set of simulated interactions, the final parameters distributions will be recorded and then used to generate the initial states of a second simulation.

The first phase consists in the positron implantation; i.e.: the process by which the high energy positron impinging onto the nanochanneled plate loses all of its initial kinetic energy to the bulk material. We will include for practicality also the process of diffusion into the material until the eventual annihilation or the conversion into Ps in case the diffusing positron reaches a material-vacuum interface.

The second phase is that of the thermalisation of the positronium during its diffusion inside of the nanoscopic channels, taking place from the instant the Ps is produced at the material interface and ending when the Ps annihilates inside of the channel or when the positron is emitted from the channel in a trajectory that does not intersect the solid surface anymore. The second phase simulation will be again split into simpler simulations but this further factorization is better explained alongside the employed model, so I'll delay its specification until section 1.5.

Splitting simulation processes has the advantage of forcing us to manipulate quantities[1] which are usually hidden within longer simulation processes; allowing us to study how these are influenced by the various simulation parameters and, ultimately, to understand more details of the physics involved in the simulated phenomenon.

1.2 Positron stopping in materials

Positrons implanted inside of a solid or liquid are slowed down to thermal energies by their interaction with the bulk material[45]. A fraction of the implanted positrons reaches the material surface again before the thermalisation process is concluded and is emitted as epithermal positrons. Thermal positrons

[1] in our case, to give an example, effective implantation profiles (defined in section 1.4.1) or interaction count spectra defined in section 1.5.3

diffuse inside of room temperature silicon for 200 nm on average. Part of the implanted positrons will annihilate during the diffusion process; most of the positrons that reach a channel surface will be converted into positronium with an emission energy of 3 eV (see section 1.5 for a more detailed discussion of the expected production energy spectrum).

We define the implantation profile $\mathcal{P}(z)$ as the probability density of an implanted positron to reach thermalisation in the interval $[\, z, z + dz \,]$ with z being the depth measured from the material surface. We'll further define the effective implantation profile $\mathcal{P}_e(z)$ as the probability density of an implanted positron to reach thermalisation and diffuse to a channel wall in the interval $[\, z, z + dz \,]$. In the first case we'll consider only positrons that actually reach thermalisation (thus ignoring epithermal and annihilating positrons) and in the second case only positrons that actually reach thermalisation and then the channel wall, therefore:

$$\int_{-\infty}^{0} \mathcal{P}(z)dz \;=\; \int_{-\infty}^{0} \mathcal{P}_e(z)dz \;=\; 1 \tag{1.1}$$

As of today no direct measurement of the implantation profile has been performed neither for bulk materials, nor for NCPs, therefore physical description of the thermalisation and transport mechanisms have to rely on theoretical models and computer simulations[58].

To describe the implantation profile $\mathcal{P}(z)$ for homogeneous solid materials a closed-formula parametrization was proposed by Makhov[59, 60, 61] and is commonly known as the Makhovian implantation profile:

$$\mathcal{P}(z) \;=\; -\frac{d}{dz} \exp\left[-\left(\frac{z}{z_0} \right)^m \right] \tag{1.2}$$

Where z is the implantation depth while m and z_0 are the parameters upon which Makhov's profile depends (the latter being, incidentally, also proportional to the mean of the $\mathcal{P}(z)$ distribution).

The Makhovian implantation profile synthesizes in a single formula a good approximation of the implantation profile and has been widely used to fit both experimental data[62, 63] and results of Monte Carlo simulations[64, 65]. By means of fits to the experimental data it was shown[65] that the mean implantation depth z_0 is mainly dependent on the density of the material and the implantation energy and that it can be approximated through the empiric formula:

$$z_0 \;=\; \frac{1}{\Gamma(1 + 1/m)} \cdot \frac{A}{\rho} \cdot \left(\frac{E}{\mathrm{keV}} \right)^n \tag{1.3}$$

where E is the implantation energy, ρ the material density, A and n the free parameters of the model. By combining the empiric formulae in equations 1.2 and 1.3 we can develop a material-independent implantation profile model featuring the values for A, m and n that best fit the experimental data across a wide variety of materials. In practice, being eq. 1.2 and 1.3 empiric formulae the best results are obtained when fitting data for materials with characteristics as close as possible to the one to which the model is applied; in our case whenever we need it we will use the values gives by Soininen[62] for silicon:

$$\begin{aligned} m &= 1.91 \\ n &= 1.69 \\ A &= 2.7\,\mu g/cm^2 \end{aligned} \tag{1.4}$$

The Makhovian profile, like most of the proposed models for implantation simulations, is formulated under the assumption of the target being composed of an uniform isotropic material.

Although the shape of the implantation profile plays a fundamental role in the design of NCPs, the best choice available as of today to predict the implantation profile is to use the Makhovian profile with an effective density. Such model has never been validated and one of the main criticism that could be moved to the blind assumption of a Makhovian profile is that the channeled portion of the NCP is a highly anisotropic material: if we draw a straight line intersecting the NCP channels and measure the length of all of its portions that traverse a void, the probability distribution for these lengths depends heavily on the orientation of the line with respect to the macroscopic surface. The fact that a positron traveling with an orientation similar to those of the channels is able to traverse large portions of void without being deflected or slowed down is, in principle, a good argument to expect the channels to provide some sort of "guidance" deepening the implantation profile further that the mere presence of void would. As we will see from the result of our simulations, this is not the case.

Two ingredients are required to compute the implantation profiles $\mathcal{P}$ and $\mathcal{P}_e$ in the NCP: a geometric model of the nanoscopic channels and a model of the interactions taking place between the material and the positrons.

1.3 Geometric model

To be able to perform the Monte Carlo simulation with objects of a specific geometry we need to be able to compute two functions that we name the *InclusionCondition* and the *NearestIntersection*. The former consists in determining, given a point in the three dimensional space, whether such point is

part of the volume of the solid object or not. The latter consists in determining, given a ray in the three dimensional space, whether that ray intersects the object's surface and, if so, the distance at which the nearest intersection lies, and the normal to the object's surface at the intersection point. For simple primitives such as cubes, spheres, cylinders and cones these functions can be written analytically. In general any solid obtained by union, intersection, subtraction or complement of the volume of the aforementioned solids (and any other analytically solvable one) can also be written analytically. Usually if some constraints are known about the solids to be combined (such as two axes being collinear) the analytic solution can be simplified and its computation eased.

1.3.1 Procedural modeling of NCPs

We will employ a geometric model of the NCP based on the composition of surfaces whose *InclusionCondition* and *NearestIntersection* are analytically expressible avoiding, thus, the necessity to approximate the surfaces with large amounts of simpler primitives (as is often done, e.g. in computer graphics or in many Monte Carlo simulations where surfaces are approximated by triangle-faced polyhedra). This will limit the models we can employ but provides a great improvement in terms of computational efficiency and stability of the results.

The shape of an NCP can be approximated as a plane surface in which a forest of irregularly-shaped channels has been carved out. The model we will be employing to describe it is based on a randomly generated geometric shape that we call *Dendrite*, whose goal is to approximate the shape of a single channel. The final shape of the NCP we will be simulating in is obtained by randomly generating numerous uniquely-shaped channels and by subtracting their volume from that of a solid semispace. The generation of a dendrite requires a random or pseudo-random number generator and six parameters: the number of segments n, the segment mean length μ_ℓ, the segment length variance σ_ℓ, the angular variance σ_θ and the maximum and minimum radii R_{max} and R_{min}. Moreover, a point on the surface of the material is needed as the origin of the dendrite.

To build a dendrite we start by building a skeleton consisting of a series of n consecutive segments starting from the origin point and concatenated so that each segments second endpoint is the first endpoint of the next segment (see figure 1.1a). Segments lengths are randomly drawn with a normal distribution with mean μ_ℓ and variance σ_ℓ. To permit the determination of the maximum length that such skeleton can have, lengths outside the range $\mu_\ell \pm m \cdot \sigma_\ell$ are rejected and drawn again, m is chosen to be large enough to make such rejection scarcely relevant. The angle between the normal to the reference plane and the first segment is drawn from a Gaussian having mean 0 and

variance σ_θ. The angle between the prolongation of each segment and the successive segment is drawn from the same distribution. Then to each segment endpoint is assigned a radius drawn from a flat distribution ranging from R_{min} to R_{max} (see figure 1.1b). Endpoints in common between successive segments share the same associated radius.

A truncated cone is built for each segment in the skeleton (see image 1.1c). Each truncated cone has the segment as axis and as base radii the radii associated to the segment's endpoints. To create a connection joint between successive cones and to smoothly cap the entire structures, a sphere is built at every endpoint using the endpoints itself as center and the endpoint's radius as radius.[2] The first truncated cone is prolonged so that its base does not intersect the reference plane. We define as the dendrite's volume the union of all of the volumes of the so-built spheric slices and truncated cones. Examples of dendrites are shown in figure 1.1d.

The NCP is then defined as a half-space from which the volume of a population of dendrites have been carved out. This construction might make the usage word *surface* ambiguous since it now can refer to the surface of the half-plane we used to build the NCP or to the surface of the final solid (which comprises also all of the channels); whenever we find that equivocation is possible we will refer to the former as the *macroscopic surface* of the NCP (being the surface seen by naked eye) and the latter as *nanoscopic surface* (being the structure visible through SEM imaging), see figure 1.2. The dendrites are all built using the same construction parameters, the surface of the half-space as the reference plane and uniformly scattered points upon such plane as origins.

When defining the dendrite geometry we introduced the forceful rejection of any segment length above $\mu_\ell + m \cdot \sigma_\ell$. This modification generated geometries that are extremely close to those generated without the rejection and allows us to determine the maximum length of a dendrite: in fact under such condition we know that a dendrite cannot not include any point whose distance from the dendrite's origin point is greater than:

$$\ell_V \;=\; n\,(\mu_\ell + m \cdot \sigma_\ell) \;+\; R_{max}$$

which, as we will see, will be extremely useful.

We divide the macroscopic surface of the NCP into squares of side ℓ_V, each identified by a pair of integer coordinates. We refer to the volume inside of the NCP identified by a cube built upon one of such squares on the non empty side as *Voxel* (see figure 1.2). The dendrites originating from the external face of each *Voxel* are generated with a pseudo-random number generator seeded with

[2] In practice, since most of the volume and surface of the employed spheres is immersed inside of the truncated cones, we employ properly cut spheric slices as primitives, operation that results in a significant computational advantage and in more numerically stable code.

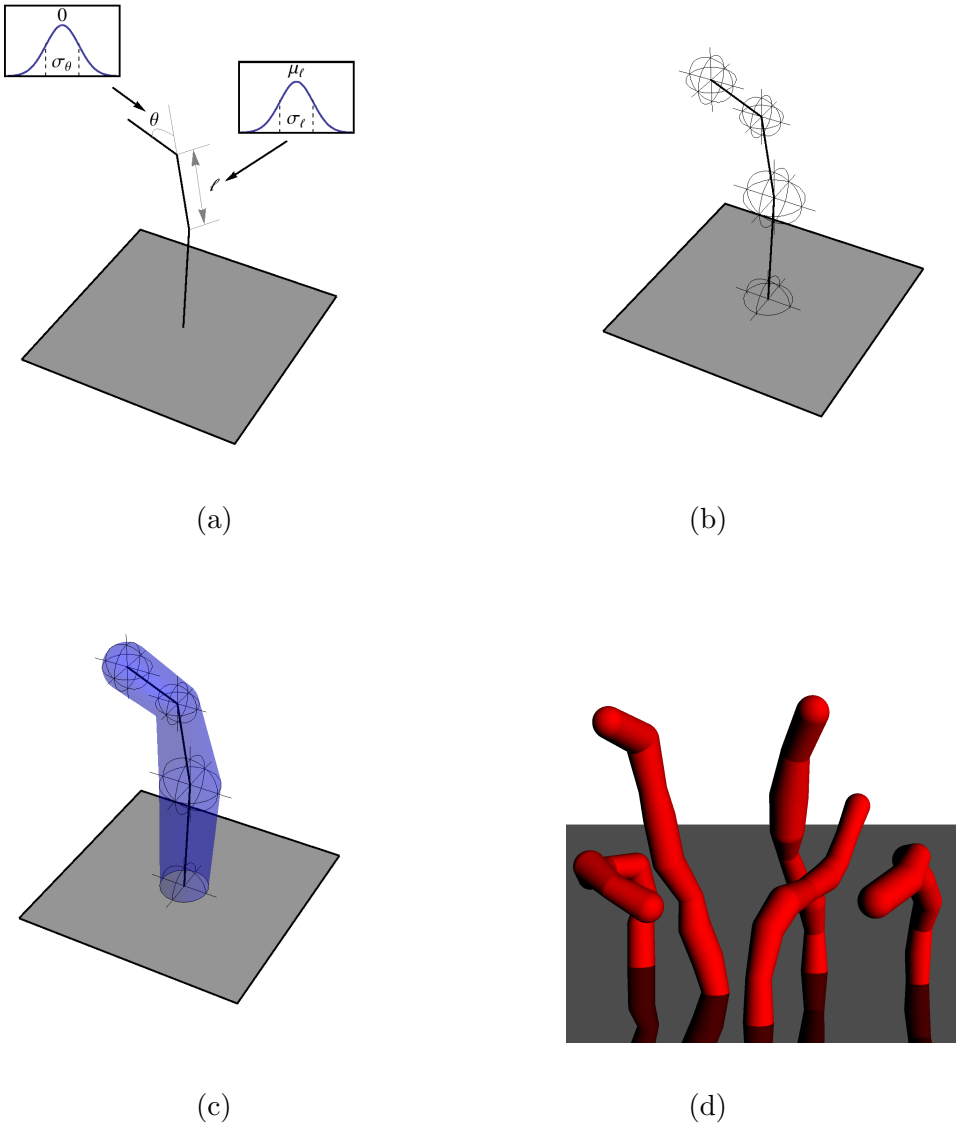


Figure 1.1: The construction of a dendrite geometry. (a) The bare skeleton of a dendrite is generated. (b) The radii are assigned to each vertex. (c) The volume of the dendrite is laid out. (d) A render of several randomly generated dendrites (parameters to generate which have been chosen to clarify their structure: dendrites used to model the NCP are thinner, longer and straighter).

the *Voxel*'s integer coordinates. Since every time the procedure is repeated the pseudo-random number generator is primed with the *Voxel*'s coordinates, it will always, in a given *Voxel*, generate the same dendrite structure, regardless of how many times the procedure is repeated. It is therefore possible to generate the dendrite population of each *Voxel* on demand thus eliminating the necessity of keeping a memory consuming description of a vast NCP.

Due to the aforementioned property of the dendrites, any dendrite intersecting a given *Voxel* must have originated from the specific *Voxel* or from one of the eight neighboring it. We call *Cluster* such set of nine *Voxels* and *Neighborhood* the set of dendrites actually intersecting the volume of a specific *Voxel*. The entire process is represented in figure 1.3. Caching systems for both *Clusters* and *Neighborhoods*, smart sorting of the tracing jobs to be executed and the addition of octrees in the construction of the *Neighborhood* descriptor objects ensure a high computational efficiency with an average time to compute a collision distance inside of an NCP of about than 50 µs for a single thread, thus configuring this kind of simulation as computationally heavy. Technical detail about how this optimizations have been put in place can be found in appendix A.3.

1.3.2 Procedural modeling of PSPs

Besides the NCP, another geometric structure has been developed to study the role of the object's geometry in the implantation process. We called it *Porous Surface Plate*, hereafter PSP. It is an object similar to the NCP in the sense that it is an half-space whose volume near the surface has been subdivided into *Voxel* structures used to generate random geometries whose volume is then removed from the half-space to obtain the final object. In the case of the PSP, however, we use spheres instead of dendrites, thus obtaining a sponge-like structure.

The reason for the introduction of this geometry is to test whether the exact geometric shape is relevant to the implantation depth (with phenomena such as the guidance descried in section 1.2) or if any geometric structure that present an analogous amount of empty space will produce similar implantation profiles. To evaluate the amount of empty space present at a specific depth we employ a parameter that we call *Occupancy*, defined as probability that a randomly chosen point at a specific depth will be part of the solid volume, thus estimating the portion of occupied volume around a specific depth.

The *Occupancy* of a PSP structure can be tuned by variating the size and amount of spherical holes; nonetheless if we want to properly reproduce also the *Occupancy* of the transition zone between the region containing the channels and the underlying solid bulk we need to make at least one of these parameters dependent on the depth, for practical reasons the holes radius is the easiest to alter. We found an empirical formula (equation 1.5, as we will

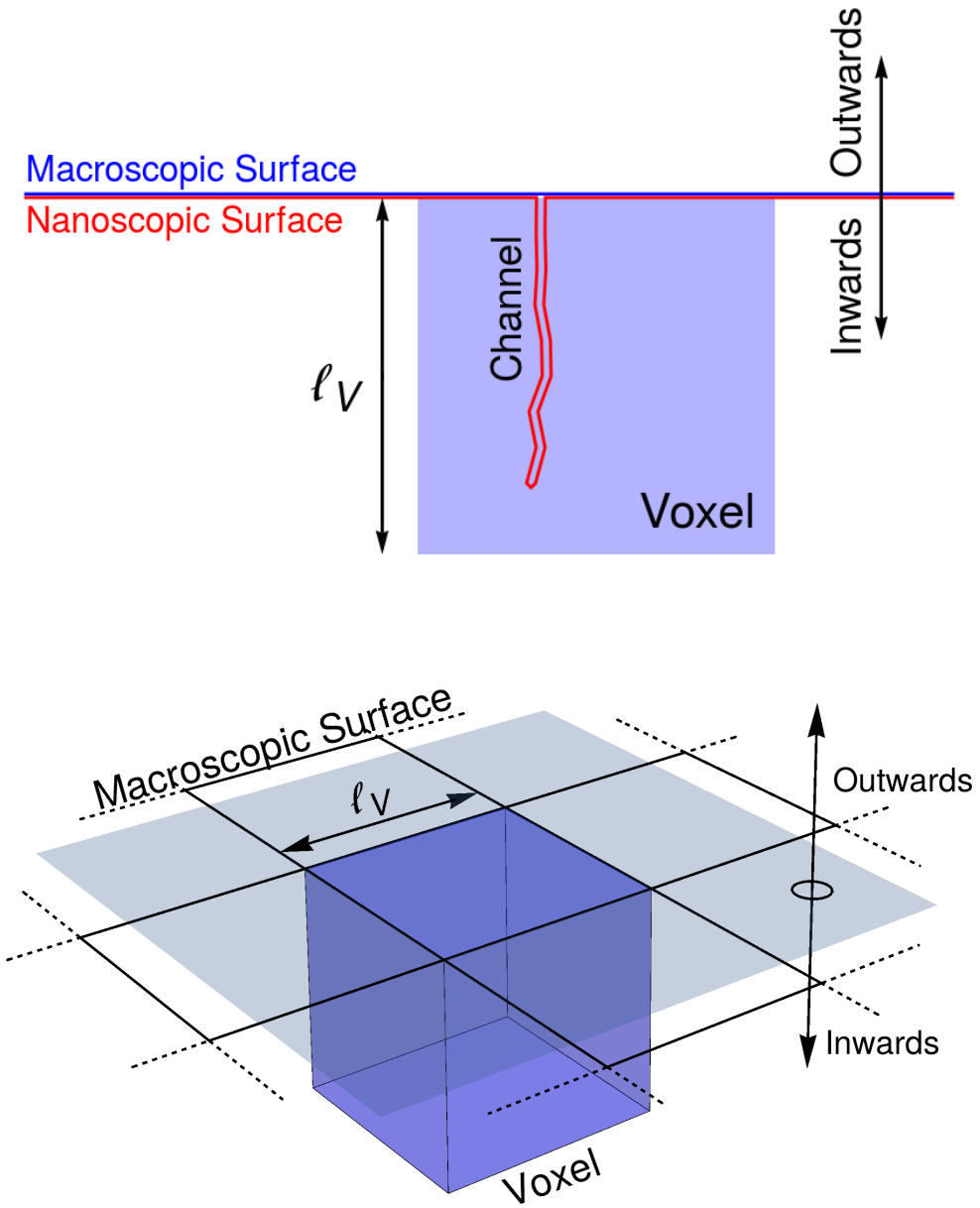
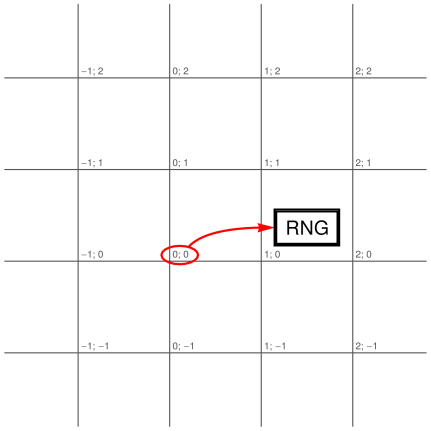
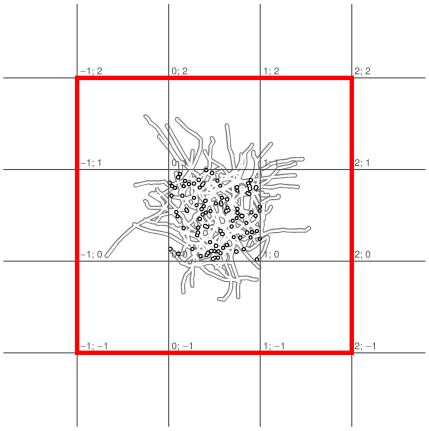


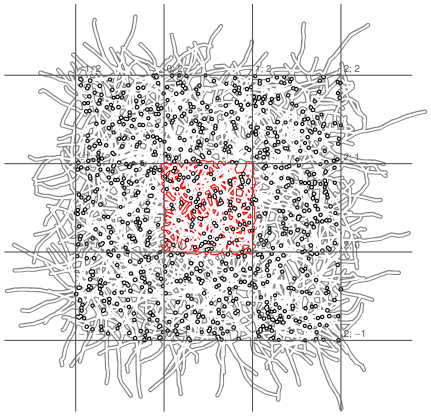
Figure 1.2: Geometric construction of the *Voxel* geometry from the macroscopic NCP surface, shown in section and in 3D projection. After carving out the dendrites we reach the definition of the nanoscopic NCP surface.



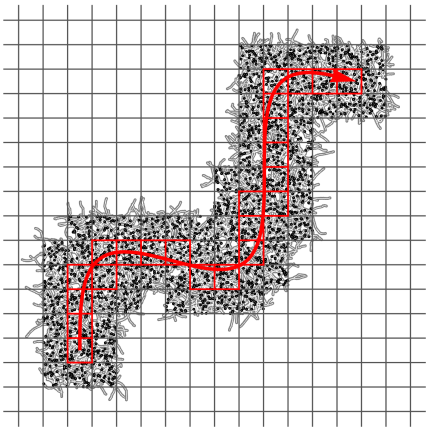
a) The macroscopic NCP surface (here seen from above) is divided into regions called *Voxels*, the integer coordinates of which are used to seed a pseudorandom number generator. This mens that every time the process is repeated each *Voxel* will always yield the exact same geometry.



b) The random number generator is then used to generate all of the channels originating from the specific *Voxel* (the aformentioned *Cluster*). *Voxel* size is chosen so that these channels cannot extend more than a *Voxel* size beyond their originating *Voxel*.



c) By computing the channels generating from 9 adjacent *Voxels* we can extract the channels (or channel portions) intersecting the central *Voxel* (what we called a *Neighborhood*).



d) During the simulation, as a particle traverses the volume occupied by the NCP, *Voxels* we can dynamically generate only the required the portion of the NCP. For clarity we exaggerated the particle's lateral motion.

Figure 1.3: The generation process for NCPs. In the illustrations the NCP is shown seen from above and, for sake of clarity, the angular variance σ_θ has been exaggerated.

see) that closely mimics the NCP *Occupancy*.

To generate the spheric holes for the PSP we employ four parameters, we call them layer thickness D, transition zone thickness J, max radius R_{max} and min Radius R_{min}. The radius of each spheric hole is randomly chosen between R_{min} and R_{max} with a uniform distribution. The depth d of the center of a spheric hole with respect to the surface is randomly chosen in the interval $[-R_{max}, D - R_{max}]$ with an uniform distribution. If $d \geq D - J/2 - R_{max}$ then the radius of the hole is multiplied by a factor ν defined as:

$$\nu \;=\; \sqrt[3]{\frac{1 \;+\; \sin\left[\,\frac{\pi}{J}(D \;-\; d \;-\; R_{max})\,\right]}{2}} \tag{1.5}$$

which is an empiric formula designed to smoothly reduce the volume of the spheric holes so that by tuning its parameters it is possible to mimic the volumetric occupancy of an NCP. The *Voxel* size ℓ_V is computed directly from the parameters as $\ell_V \;=\; D + R_{max} + J/2$. Occupancy profiles for an NCP and a PSP tuned to mimic it are shown in figure 1.4.

1.4 Implantation profiles

To be able to compute the implantation profile in NCPs we need to be able to simulate the interaction between low energy $\leq 30\,\text{keV}$ positrons with matter. We are interested only in the positron's behavior; due to the chosen energy range forbidding the secondary particles to produce any additional positrons we can ignore any other particle produced by an interaction besides the main positron.

Our interaction model comprises four interactions: *Annihilation*, *Bremsstrahlung*, *Ionisation* and *Elastic Scattering*. To be able to simulate any of these interaction two elements are required. The first is the ability to compute the average distance between successive interactions inside a specific material given the material characteristics and the particle energy. The second is a procedure to generate, given the same informations about the particle and the material, a scattering angle and energy drawn randomly so that their combined probability distribution reflects the desired physic model for the interaction.

For the first three interactions, we derived empiric formulae from the Penelope simulation code[66, 67] included in Geant4[68]. The original Geant code has been integrated inside of the simulator to implement the bremsstrahlung and ionization scattering processes. The annihilation process has been implemented as a simple end-of-track positron death. Average interaction distances computed with the Geant code for a range of materials and energies

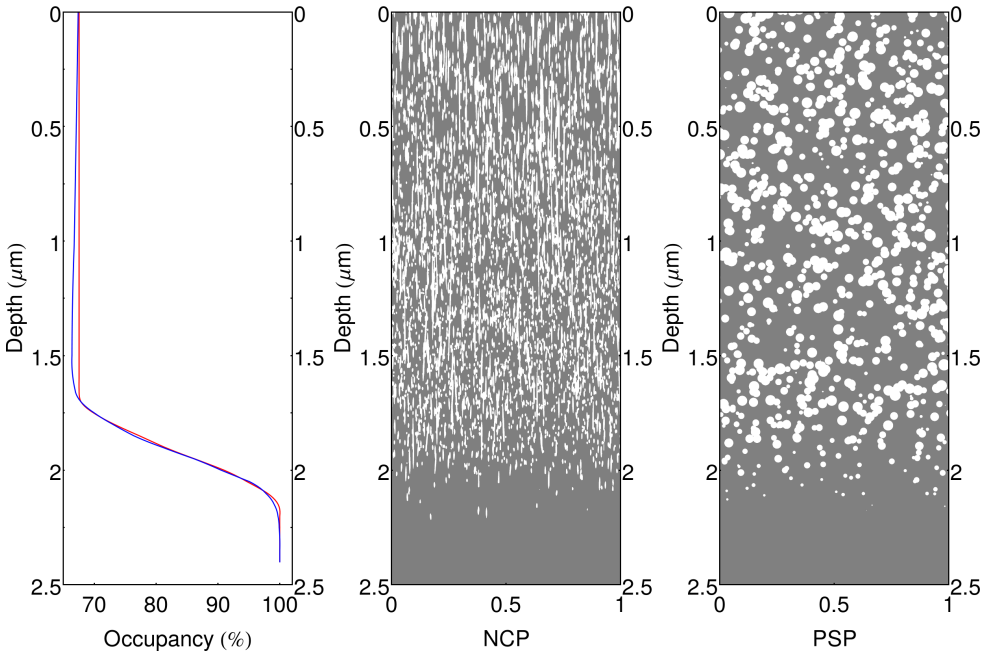


Figure 1.4: On the left the occupancy parameter as a function of the depth in the material for a typical NCP (in blue) and a PSP tuned to mimic its occupancy (in red). On the right, sharing the same y scale, a section of such NCP (in center) and of such PSP (on the right), white areas are empty spaces, gray areas are bulk material.

within the considered energy have been fitted with empiric formulae to permit a standalone implementation not requiring the Geant infrastructure.

To improve computational performance, elastic scattering on nuclei is often implemented with an effective interaction model that condenses multiple Coulomb scatterings into a single interaction[69, 70]; this is the case, also, of the Penelope implementation[67]. In our case this kind of implementation may result in artifacts since the resolution of the details in the employed geometry are comparable to the scale of the elastic scattering mean free path, the former being in the scale of the nanometers and the latter being comprised, for silicon, between 0.2 and 230 nm in the considered energy range. We implemented the elastic scattering using the model proposed by Fernandez[71]. Interaction distance and scattering angle distribution have been implemented as proposed in the article; the scattering energy is computed from the scattering angle using conservation of energy and momentum.

When during the deceleration process the positron reaches the energy cutoff threshold of 100 eV we consider the implantation process to be concluded, being the residual mean stopping path according to the Makhovian model

shorter than 1 nm.

We verified our interaction model and implementation by comparing implantation processes in solid material with implantation simulation present in the literature[58, 72, 62, 73]. Implantation inside silicon and aluminum bulks have been run with a 10 keV positron beam perpendicular to the surface. The implantation profiles obtained with our code are in good accordance with the profiles simulated with Geant 4 already present in literature as can be seen in figure 1.5.

We ran several simulations mimicking the shape of the NCPs and beams we employ in our experiments. Our workhorse NCP is described in [54] where it is named sample #0. A good choice of parameters to simulate it are assuming a surface density of channels of $10^{16}\mathrm{m}^{-2}$, $\mu_\ell = 250\,\mathrm{nm}$, $\sigma_\ell = 50\,\mathrm{nm}$, $R_{Min} = 3\,\mathrm{nm}$, $R_{Max} = 4\,\mathrm{nm}$, $\sigma_\theta = 0.3\,\mathrm{rad}$ and 8 segments per channel. A section of the geometry generated with such parameters is shown in the central panel of figure 1.4. All of the simulations presented in this chapter have been run employing the aforementioned parameters. Another NCP has been tested in the experiment described in chapter 2, namely a specimen of what in [54] is named sample #2 that is well described by the parameters of sample #0 with the exception of R_{Min} and R_{Max} that need to be edited to $R_{Min} = 5\,\mathrm{nm}$, $R_{Max} = 8\,\mathrm{nm}$. Whenever during chapter 2 we will need to compare experimental data coming from this sample with simulated data, this alternate set of parameters will be employed.

To mimic the occupancy of our model NCP we employed a PSP with volumetric density of holes $\rho = 1.1 \cdot 10^{22}\mathrm{m}^{-3}$, $R_{Min} = 15\,\mathrm{nm}$, $R_{Max} = 25\,\mathrm{nm}$, $D = 1905\,\mathrm{nm}$ and $J = 545\,\mathrm{nm}$; a section of the geometry generated with such parameters is displayed in the rightmost panel of figure 1.4. In such objects we implanted monochromatic beams of positrons having energies in the range $1 \div 30\mathrm{keV}$ and with an incidence angle relative to the normal of the NCP/PSP surface randomly drawn from a normal distribution having null mean and a standard deviation of 2°[3]. The resulting implantation profiles are shown in figure 1.6.

Our simulations show that in the the experimental conditions employed the effects of the channel geometry affect only shallow implantations (those executed with a positron beam with energy less than 3 keV). For higher energies the channeling effects expected from the NCP geometry average out and the nanoscopic structure effects boil down to those expected from a reduction of the material's density. In table 1.1 is reported the fit operated with Makhov's model (see equations 1.2 and 1.3) to the implantation profiles $\mathcal{P}(z)$ in NCPs and in bulk silicon. We fixed the n parameter since we have a limited range in which the apparent Makhovian profile is not distorted by

[3] We chose the parameters to simulate the beam of the apparatus[38] which will be the focus of chapter 2.

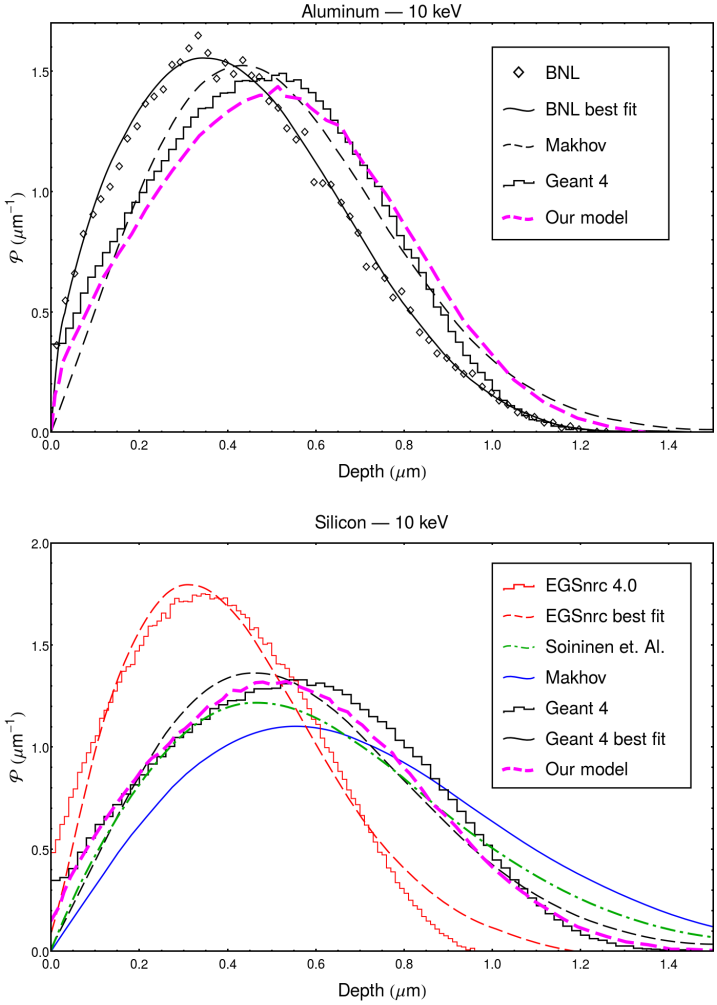


Figure 1.5: Upon both implantation profile plots taken from the work of Dryzek[58] the implantation profile obtained through our simulator has been superimposed in thick dashed magenta. The topmost plot compares implantation depth in bulk aluminum, the bottom one compares implantation profiles in bulk silicon. The solid histogram in both plot has been obtained through simulation with the GEANT4 code. The BNL code [72] simulation and its bet Makhovian fit is shown; alongside the predicted profile according to the model by Soininen et al.[62]. In the second plot the profile obtained though the EGSncr4 code[73] and its best Makhovian fit is shown. In both cases the material-independent Makhovian profile is reported. It can be seen in both plots that our implantation profile is in good accordance with Geant4 simulation.

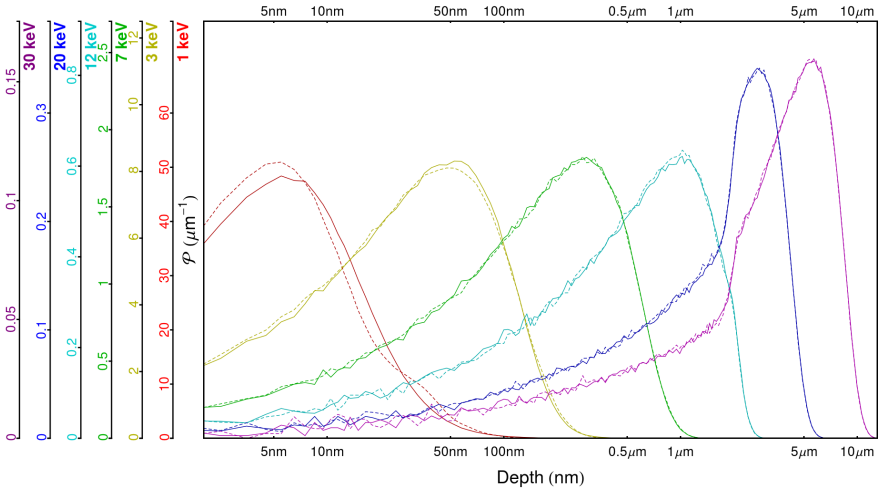


Figure 1.6: Implantation profiles $\mathcal{P}(z)$ in NCPs (continuous lines) and in mimicking PSPs (dashed lines). Color differentiate between implantation energies. The vertical scale has been chosen so that in an uniform material the implantation profiles would have the same peaks.

the nanoscopic structure (as a reference only two of the 6 profiles shown in figure 1.6 are Makhovian), therefore we cannot effectively fit the dependency on energy. The parameter A fitted over the bulk profiles corresponds to the A parameter fitted over the NCP implantation profiles if properly scaled (i.e.: if divided by the occupancy of the NCP).

1.4.1 Effective implantation profiles

To compute the effective implantation profile we need to be able to simulate the behavior of positrons below the 100 eV limit of validity of the Penelope model that we employed for the implantation. To do so we need to in introduce two additional interactions which we called *Transition* and *Diffusion*. A schematic representation of the different processes employed in each simulation phase can be found in figure 1.7.

The *Transition* interaction takes care of simulating the positron behavior from the moment it reaches the threshold of 100 eV down to the moment in which it crosses the 1 eV threshold. According to the material independent Makhovian model (see section 1.2) we expect a 100 eV positron to travel about 0.37 nm in silicon which means that it can be probably ignored completely without altering the final result. Out of conscience we preferred to include an empiric model also for the transition interaction. We set it up with a constant mean interaction distance $\lambda = 0.02$ nm and generating interaction processes

Implantation energy	n (bulk)	m (bulk)	A (bulk)	$A/0.68$ (bulk)	m (NCP)	A (NCP)
1 keV	1.82	1.71	1.76 µg/cm^2	2.55 µg/cm^2	1.42	2.59 µg/cm^2
3 keV	1.82	1.71	1.76 µg/cm^2	2.55 µg/cm^2	1.57	2.22 µg/cm^2
7 keV	1.82	1.71	1.76 µg/cm^2	2.55 µg/cm^2	1.71	2.28 µg/cm^2
12 keV	1.82	1.71	1.76 µg/cm^2	2.55 µg/cm^2	1.95	2.67 µg/cm^2

Table 1.1: Result of the fit to the implantation profiles in bulk silicon and in the NCP geometries (see figure 1.6) using the Makhov model (see equations 1.2 and 1.3). In the fifth column the parameter A obtained from m fitting the implantation profile in bulk silicon divided by the NCP occupancy at a depth of 0.5 µm, we found that the scaled A parameter is in fair accordance with the NCP-fitted A, indication that implantation in the nanochanneled region can be well approximated with a solid bulk having a reduced density.

Figure 1.7: Different processes used in the implantation. From left to right: i) The implantation profile is computed using the interactions obtained from the *Penelope* model ii) The particle is further propagated through an effective interaction until it reaches an energy of 1 eV iii) The diffusion process is simulated as a series of discrete interactions until the particle either annihilates or it reaches a channel wall.

that do not deflect the particle from its original trajectory but decrease its energy after every single interaction by a constant fraction $f = 0.25$. With these parameters it takes at most 17 transition interactions for the positron to reach 1 eV and during the transition process the particle travels on average 0.34 nm which is compatible with the value expected from Makhov's model.

An alternative to Makhov's model to compute the range of a 100 eV positron in silicon before it reaches thermal energies is the model proposed by Nieminen and Oliva[74]. According to this model we can estimate the range of a positron in aluminum before it reaches temperatures at which it cannot be treated as a free particle undergoing electron scattering (that Nieminen and Oliva set to 5 eV), we obtain 17 nm. If we decided this to be a fair estimate for the range of positrons in silicon, then the effective interaction described in the previous paragraph would need to be differently calibrated; as previously stated for the computations we will employ our model the choice between these two alternatives is utterly irrelevant.

The *Diffusion* interaction kicks in below 1 eV and needs to model the brownian motion of thermal free positrons in the material. We simulate the positron diffusion as a series of random scatterings after each one of which a new direction for the positron motion is selected isotropically and a new momentum is randomly drawn from a thermal distribution having the temperature of the material the positron has been implanted in. The mean interaction distance for the diffusion process has been computed as [75]:

$$\lambda_{diff} = D \cdot \sqrt{\frac{3m^*}{k_B \cdot T}} \tag{1.6}$$

with D the positron diffusion coefficient in the selected material, k_B the Boltzmann constant, T the material temperature and m the effective positron mass, good choice for which is 1.5 times the positron mass. D can be computed as [75]:

$$D = \tau_r \frac{k_B T}{m^*} \tag{1.7}$$

where τ_r is the relaxation time for the dominant scattering mechanism, according to [45] $\tau_r = 7.2 \cdot 10^{-14}$s.

During the *Transition* and the *Diffusion* phases the *Ionization*, *Bremsstrahlung* and *ElasticScattering* interactions are disabled with the new interactions taking their place in the simulation process. Still the new interactions cannot simulate the effects of positrons annihilations, so we will need to extend the annihilation interaction below the 100 eV threshold. By fitting the Penelope model we found out that in the range 100 eV $\rightarrow$ 30 keV the mean free path for the annihilation process can be well approximated by:

$$\lambda_{ann} \;=\; A \cdot \sqrt{E} \tag{1.8}$$

with E being the positron's energy and A a fit parameter. For silicon we found $A = 113.5\,\mu\mathrm{m}/\sqrt{\mathrm{eV}}$. This model can be naïvely extended beyond its original formulation boundaries. From Monte Carlo simulations we see that this naïve extension yields a diffusion length for room temperature silicon of 248 nm which is surprisingly close to the experimental value of 219 nm given by Schultz and Lynn[45]. We nonetheless aim to a higher degree of accuracy so although we will accept the naïve extension for the *Transition* energy range we modified the annihilation mean free path to reflect experimental values for the positron diffusion length in silicon.

The positron diffusion in silicon can be written as[75]:

$$L \;=\; \sqrt{\tau_{ann} D} \tag{1.9}$$

with τ_{ann} the mean positron lifetime in the chosen material. In room temperature silicon $L \approx 200\,\mathrm{nm}$. Due to the dependency (see eq. 1.7) of the diffusion coefficient on the material temperature, $L \propto \sqrt{T}$. Let's now consider a diffusion process terminating in an annihilation. The average number of diffusion interactions in the process can be written as the annihilation mean free path divided by the diffusion mean free path:

$$\langle\, N \,\rangle \;=\; \frac{\lambda_{ann}}{\lambda_{diff}} \tag{1.10}$$

Let's compute the average distance from the origin that a particle undergoing a random walk with exponentially drawn step length travels after N steps. After a single step the probability distribution for the final position is given by:

$$\mathcal{W}(\vec{x}) \;=\; \frac{1}{4\pi\lambda_{diff}|\vec{x}|^2} \cdot \exp\left[-\frac{|\vec{x}|}{\lambda_{diff}}\right] \tag{1.11}$$

This is probability distribution with mean 0 and variance:

$$\sigma^2_{\mathcal{W}} \;=\; \frac{2}{3}\,\lambda^2_{diff} \tag{1.12}$$

When considering the position $\vec{x}_N$ of the particle after N scatterings, we know that for large N in virtue of the central limit theorem $\vec{x}_N$ has a normal distribution with zero mean and variances:

$$\langle x_N^2 \rangle \;=\; \langle y_N^2 \rangle \;=\; \langle z_N^2 \rangle = \frac{2}{3} N \lambda_{diff}^2 \tag{1.13}$$

The distance from the origin is then $L = \sqrt{x_N^2 + y_N^2 + z_N^2}$ and being x_N, y_N, z_N independent normal random variables L follows a χ^2 distribution having mean:

$$L \;=\; \sqrt{\frac{4}{3}} \cdot \frac{\Gamma(2)}{\Gamma\left(\frac{3}{2}\right)} \lambda \sqrt{N} \;=\; \lambda \sqrt{\frac{16\,N}{3\pi}} \tag{1.14}$$

Which is proportional to $\sqrt{N}$ and to the average step length. Therefore:

$$L \;\propto\; \lambda_{diff} \cdot \sqrt{\frac{\lambda_{ann}}{\lambda_{diff}}} \;\propto\; \sqrt{T} \tag{1.15}$$

$$\Rightarrow \quad \lambda_{ann} \;\propto\; \frac{T}{\lambda_{diff}} \tag{1.16}$$

Taking into account also the dependency of the mean free path in the temperature (see eq. 1.6 and 1.7):

$$\lambda_{ann} \;\propto\; \sqrt{T} \tag{1.17}$$

we can obtain the proportionality coefficient by imposing the experimental value for the diffusion length given by Schultz and Lynn[45]. We determined that:

$$\lambda_{ann} \;=\; \sqrt{T \cdot 373\,\mathrm{fm/K}} \tag{1.18}$$

yields correct diffusion lengths.

We computed the effective implantation profiles as defined in section 1.4 for samples held at temperatures of 20 K and 300 K both for the NCP and the PSP geometry. By comparing the four resulting profiles we found that both the change in temperature and the change in geometry affect the shape of the resulting spectra; nonetheless the effects of the change in geometry are dramatically more prominent than the effects induced by the change in temperature (see figure 1.8a). Indeed as was the case with implantation profiles in section 1.4 when implanting at energies greater than 3 keV varying the geometry does not affect the overall shape of the effective implantation profiles as can be seen from figure 1.8b.

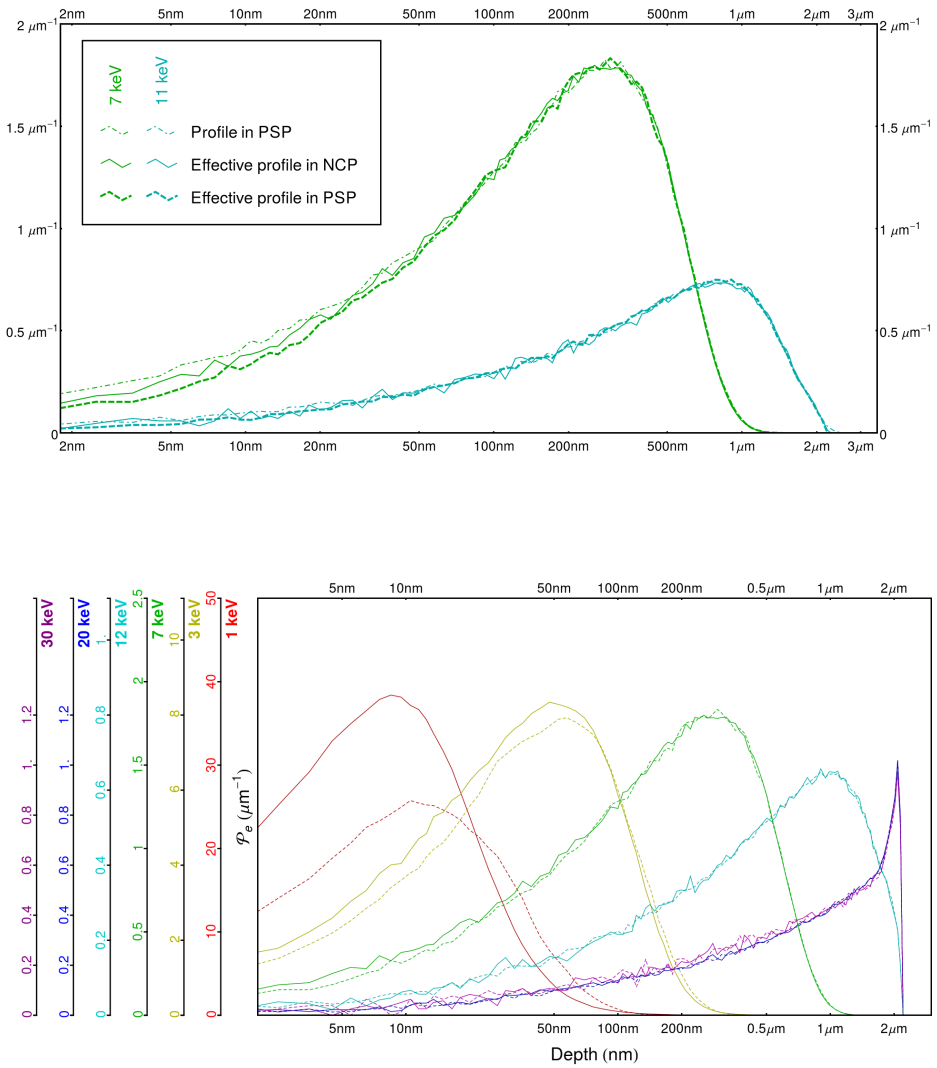


Figure 1.8: Above: comparison of effective implantation profiles in NCP and PSP with the simple implantation profile in NCP for 7 keV and 12 keV implantation energy and the material temperature set to 20 K. Below: effective implantation profiles in NCPs (continuous lines) and in mimicking PSPs (dashed lines). Color differentiate between implantation energies.

We can confidently conclude, therefore, that effective implantation profiles simulated with different geometries (NCP/PSP) or different temperatures within the $20 \div 300$K range can be used indifferently as input for the next simulation step. Instead if different implantation energies or different occupancies (see section 1.3.2) are employed, specific simulations need to be run to determine the appropriate implantation profiles. Being the NCP geometry extremely more computationally costly to employ the possibility to switch to the PSP geometry to simulate effective implantation profiles provides a huge computational advantage that can be spent either in the form of reduction of computation time or in an increase of the final result precision.

We can also observe that for implantation energies higher than 3 keV and lower than 12 keV the implantation profile is almost identical to the effective implantation profile due to the dense network of channels creating a trapping-like phenomenon[76, 77, 78]. This means that within this energy range we can avoid the diffusion simulation step altogether to further increase the computational efficiency.

Another parameter we need to extract from the implantation simulation is the surface efficiency S defined as the probability that a positron implanted into the sample will thermalize and reach a channel's surface before annihilation. This parameter will be essential in section 1.6 to estimate the annihilation probability during the positronium thermalisation inside of the channels. Surface efficiencies for NCP and PSP geometries are shown in figure 1.9 plotted as a function of energy; as expected the surface efficiency is highly dependent on the employed geometry.

1.5 A thermalisation model

We divided the simulation process into two phases. The goal of the second of which is to propagate a newly-formed positronium atom along the channel, letting it interact with the channel's inner walls and lose energy up until the point it annihilates in the sample or it is re-emitted from a pore.

To perform this second operation we are not bound to employ necessarily the same geometry used to implant positrons: in fact it is interesting to see whether employing finite-length straight cylindrical channels or single dendrites from the NCP structure has any effect on the results of the thermalisation process. Whenever a straight cylindrical channel or a single dendrite can be used in lieu of the complex NCP geometry, their employment will have the advantage of requiring less computational power and, in the case of the cylindrical channel, allowing for the analytical resolution of some problems (see, for example, theorem 1 in section 1.5.2). In all of the following passages we studied all three cases and whenever of interest reported the comparison of the results obtained from different geometries.

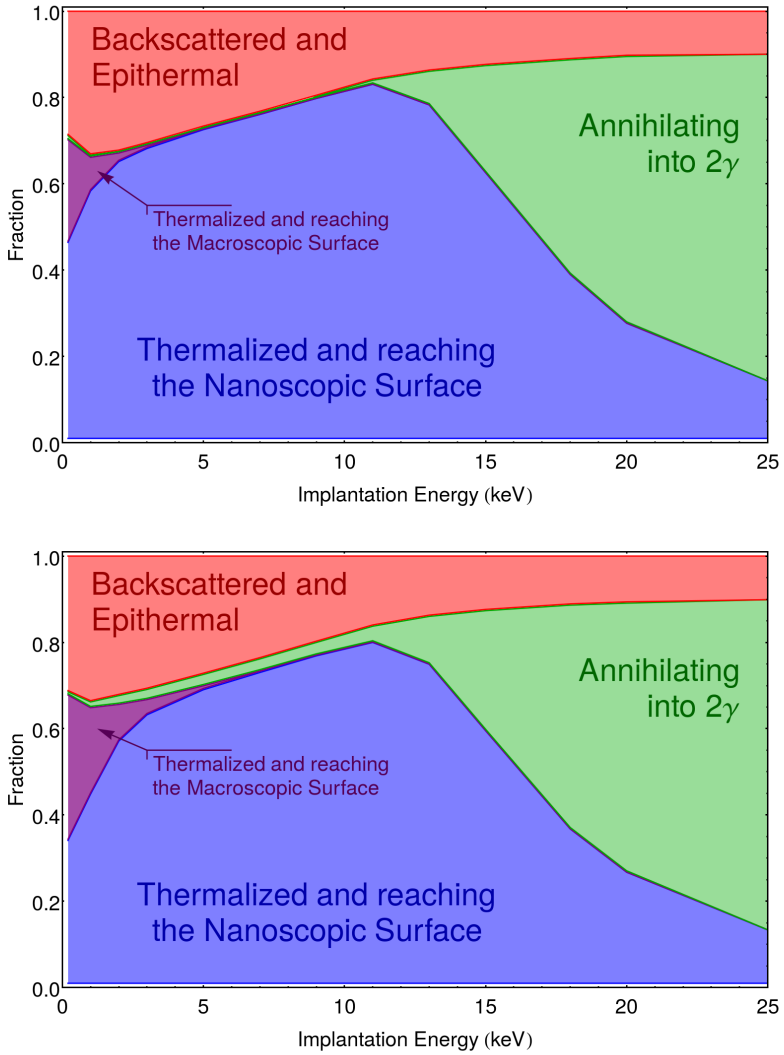


Figure 1.9: Fraction of the implanted positrons reaching different final states after implantation and diffusion in the NCP geometry (above) and in the PSP geometry (below). We differentiate between positrons that reach the macroscopic surface with an energy above 1 eV (Red), positrons that reach it with energies below 1 eV (Purple), positrons that reach the microscopic surface with an energy below 1 eV (Blue) and positrons that annihilate in the material bulk (Green). The upper edge of the blue region is the function defined as surface efficiency at the end of section 1.4.1; its sudden drop when the implantation energy crossed the 15 keV threshold is due to the implantation distribution exceeding the length of the nanoscopic channels.

To do so we will again use the Monte Carlo technique by modeling the interactions between the positronium and the channel's walls as a series of discrete interactions. The validity of this approach is hardly disputable shortly after the positronium has been formed, due to the Ps production energy (3 eV) giving the newly-formed atom a De Broglie wavelength that is much smaller than the scale of the channel's geometry. Nonetheless, as we approach the end of the thermalisation process and the energy of the positron approaches the thermal energy scale, its De Broglie wavelength becomes comparable to the radius of the nanoscopic channel; when this happens the discrete interaction model might lose its validity[57, 56]. No effective potential to model this behavior has yet been proposed and performing a complete quantum simulation of the interactions between the positronium wavefunction and the channel walls requires advanced computational tools and a great deal of computational power, both going beyond the scope of our simulator. We'll therefore proceed to implement a classical discrete-interaction simulation of the thermalisation process and test it against the available experimental data.

The energy spectrum of positronium produced from SiO_2 has been measured by Nagashima et al. in 1998 [79], who showed it to manifest two main emission energy peaks located at 1 eV and 3 eV. The authors attributed these two peaks to two different production mechanisms, namely the production of pseudo-Ps in the material bulk (1 eV) and the formation of positronium at the material-vacuum interface (3 eV), consideration supported by the observation made by Paulin and Ambrosino in 1968[80] highlighting the dependency on the positronium production efficiency from silica powders on the size of the powder grains. Based on the data from [80] we can say that, due to our sample's SiO_2 coating being too thin to provide a significant pseudo-Ps yield, we can assume a δ distribution centered at 3 eV as a suitable production spectrum for positronium inside of the sample's channels.

We'll model the process of scattering of a positronium atom against a nanochannel's wall using a single interaction process. This process needs to determine the scattering angle and the new energy of a positronium after it has collided with the wall, provided the energy of the impinging positronium, its incidence angle, the local wall curvature and the material temperature. Written in its full extent the model is a function that takes the aforementioned parameters and associates to them a three-dimensional distribution in the phase space of outgoing positronium momenta (1 dimension) tensor the outgoing positronium scattering angle (2 dimensions).

Since as of today no experimental data is available inquiring directly the scattering interaction between a positron and a SiO_2 coated surface we'll try to formulate the simplest reasonable model we can to describe it. We'll start by drawing in some common assumptions to reduce its complexity. First of all we will assume completely isotropic scattering angles, i.e.: we'll assume that the outgoing positronium angle is uniformly distributed in the 2π solid angle

allowed by the surface geometry. This model is similar to the one used to simulate the interaction of light with corrugated surfaces, like sheets of paper of waterpainted walls, and is antithetic to a completely correlated model, as it could be the one used to scatter photons upon polished or glossy surfaces. Our choice is motivated by the fact that SEM inquiry of nanochanneled plates showed extremely irregular surfaces that can degenerate, in some cases, in additional chambers sprouting out of the main channel's shaft.

The second assumption that we will make is that the final positronium energy is independent of its scattering angle with the surface. The reason of this is that, due to the irregular shape of the surface, we expect the scattering angle to be mostly due to the irregularity of the wall and not to the physics of the interaction. Moreover we'll assume that, whatever model we'll decide to employ to compute post-interaction energies, it will have to asymptotically tend to a thermal distribution having the sample's temperature. In physical terms this translates to require that infinitely lived positronium trapped in an infinitely long channel will eventually reach thermal equilibrium with the channel.

This final condition is actually quite restrictive. A model that respects it can be obtained by drawing a random particle from a thermalized ideal gas, simulating an elastic scattering between the impinging positronium and such particle and taking the final positronium energy as its energy after it hit the wall. This model is extremely similar to the one proposed by Sauder[81] to describe positronium thermalisation in rare gases which has been widely and effectively used[82, 83]. In his model Sauder assumed isotropic elastic scatterings between positronium and atoms of the gaseous moderator, then computed the average energy loss after each interaction. Since we do not aim to write the average energy after n interactions but, instead, the complete energy probability distribution after n interactions we need to expand on Sauder's model and employ the entire energy-loss distribution (not only its mean) to evolve the positronium energy spectrum.

Under this premises we can split the simulation effort in three phases:

- The first phase consists in computing the energy spectrum $\mathcal{K}_n(E)$ of a distribution of Ps that has been generated at 3 eV and each member of which has undergone exactly n interactions. We'll discuss it in section 1.5.1.

- The second phase consists in computing the probability $F_S(n)$ that a positronium will be able to undergo at least n interactions with the channel walls without decaying as a result of these interactions or due to self-annihilation in flight. This is discussed in section 1.5.2.

- The third phase consists in computing the interaction count spectrum $\mathcal{S}(n)$, defined as the probability that a positronium generated on the

nanoscopic surface of a channel will interact exactly n times with the channel walls before escaping into the vacuum, assuming the positronium to be infinitely lived and the interactions with the channel walls incapable of causing its annihilation. The parameter $\mathcal{S}(n)$ condenses all of the information relative to the geometric aspect of the simulation; we'll discuss its computation in section 1.5.3.

The first and third phase can be computed independently since no correlation exists in our model between the geometric and the energetic component of the thermalisation process. Moreover their results depend on different parameters: the output of the first phase, $\mathcal{K}(n)$, depends on the effective mass M and on the sample temperature T, the output of the third phase $\mathcal{S}(n)$ depends on the implantation energy and on the channel geometry. The second phase makes use of results coming from both the first and third phase so, chronologically, it should actually be performed as the last one; we found it nonetheless clearer to illustrate it as second. The result of the second phase $F_S(n)$ depends potentially on all of the input parameter, still its influence on the final result is less dramatic than that of $\mathcal{K}(n)$ and $\mathcal{S}(n)$ so the main dependency of the final spectrum on the input parameters is given by the computation of phase one and phase three.

After all three components have been computed we can write the energy distribution $\mathcal{K}(E)$ of the emitted positronium as:

$$\mathcal{K}(E) \;=\; \frac{1}{k} \cdot \sum_{n=0}^{\infty} \mathcal{S}(n)\mathcal{K}_n(E)F_S(n) \tag{1.19}$$

with the normalization k equal to:

$$k \;=\; \sum_{n=0}^{\infty} \mathcal{S}(n) \cdot F_S(n) \tag{1.20}$$

the summations to ∞ can be cut to a finite value of n and still provide a good estimation of the limit value due to the fact that both $\mathcal{S}(n)$ $F_S(n)$ asymptotically tend to zero as an exponential law or better.

1.5.1 Phase 1: Thermalisation dynamics

If the positronium is taken to be traveling with speed $\vec{v_i}$ before it interacts elastically with an atom of mass M traveling at speed $\vec{v}$, imposing the conservation of momentum an energy does not determine the final positronium speed v_f but forces it to lay in the locus:

$$
\begin{aligned}
\vec{v}_f &= \vec{q} + \hat{u} \cdot r \\
\text{with} \quad \vec{q} &= \vec{v}_i \cdot \frac{m}{m+M} + \vec{v} \cdot \frac{M}{m+M} \\
\text{and} \quad r &= |\,\vec{q} - \vec{v}_i\,|
\end{aligned}
\tag{1.21}
$$

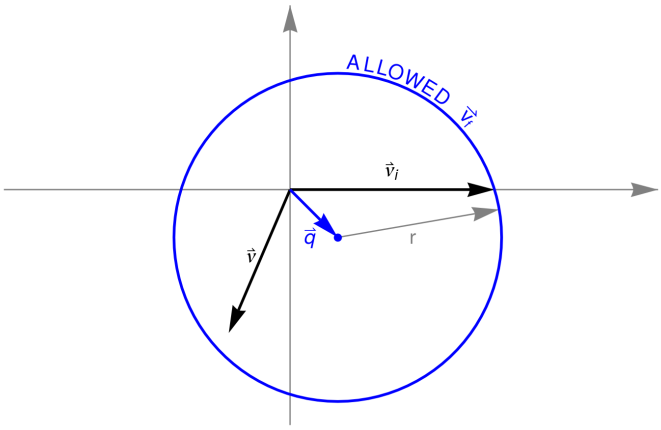


Figure 1.10: Allowed final energies for an interacting positronium. For illustration purposes, the effective mass M of the interacting particle (atom in the wall) has been taken to be twice the positronium mass, as to make $\vec{q}$, the velocity of the center of mass, distinguishable from th origin. Still even in the real world scenario in which $\vec{q}$ is extremely small, if we approximate $\vec{q} \approx 0$ no thermalisation takes place.

where m is the positronium mass and $\hat{u}$ is a random vector of length 1, and all of the vectors lay in a one, two or three dimensional space. The fastest way to reach the result shown in equation 1.21 is to begin to analyze the case of two-body decaying of a particle in the particle's center of mass reference system (which yields quickly the shape of the locus), then to change reference frame and add the second particle's momentum and energy to the system initial conditions; the complete procedure to do so is shown in detail in §16 and §17 of [84].

If $\hat{u}$ is uniformly drawn, as required by the isotropic scattering model of [81], then an arbitrary distribution of Ps will converge after a suitable number of wall interactions to a thermal distribution with a dynamic that is independent on the dimensionality in which the problem has been solved (that is regardless of $\hat{u}$ having been picked from the unitary circle, the unitary sphere or the ± 1 set). The thermalisation dynamic with M set to the mass of a silicon atom is shown in figure 1.11.

Due to the lighter element in the sample composition being the oxygen we do not expect the value of the effective mass to be less than 16 amu. Previous works that employed Sauder's model to study positronium thermalisation with silica powders[85] found values of M ranging up to 112 amu; we expect therefore realistic values for M to be comprised in the 16 amu $< M <$ 112 amu range.

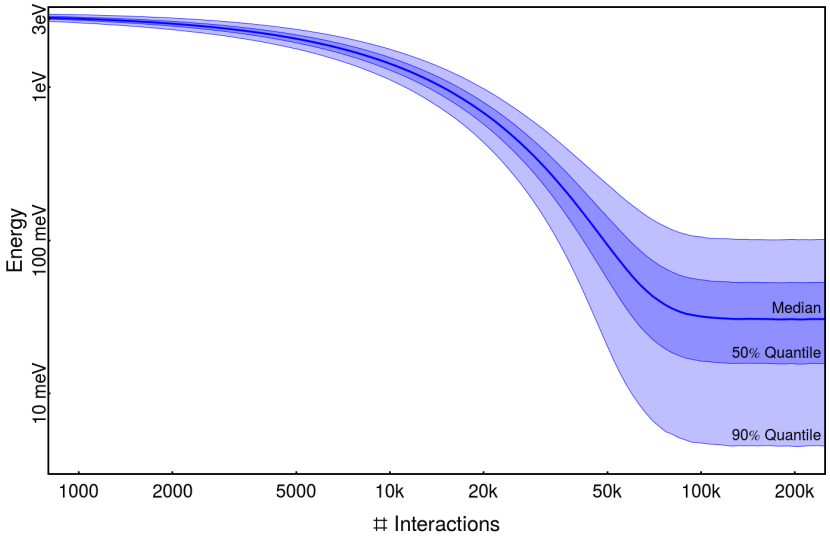


Figure 1.11: Evolution of a δ distribution obtained through Monte Carlo simulation of positrons (centered at 3 eV) after each element of which has undergone n elastic scattering against a gas of particles with a temperature of 300 K and mass equal to that of silicon.

1.5.2 Phase 2: Annihilation within the channels

We need to consider that, during the thermalisation process, positronium can annihilate inside of the nanoscopic channels; the phenomenon can be independently caused by two different processes: the positronium self-annihilation and 2γ annihilations caused by the interaction with the channel walls.

We will adopt a model similar to the one proposed in [86]: every time a positronium atom hits the wall of a channel it will have a fixed probability p of annihilating into two γs. At the same time during the whole time the positronium atom is traveling inside of the channel it is also subject to the possibility of self-annihilation with the usual characteristic lifetime $\tau = 142$ ns.

Before proceeding further let's demonstrate a quick theorem that will come in handy shortly.

Theorem 1: Let there be a cylinder with diameter 1, a point P on its surface. Let $\bar{r}$ be the radius connecting the point P with the axis of the

cylinder. Consider a polar coordinate system centered in P and having its azimuth along $\bar{r}$ in the direction pointing inwards in the cylinder. Let θ be the azimuthal and ϕ the polar angle of such coordinate system. Let's then consider a random direction drawn using a uniform distribution for the angle ϕ and an arbitrary distribution for the angle θ with the only requirement that $\theta < \pi/2$. Let $\bar{R}$ be the line directed along such direction that passes through P. Let P and Q be the intersections of R with the cylinder surface.

Thesis: The average distance between P and Q is 1.

Consider an infinite cylinder with unitary radius and its axis coinciding with the $\hat{z}$ axis of a Cartesian coordinate system. With no loss of generality let's place the point $\vec{P} = \{0, -1/2, 0\}$ (see figure 1.12). Let's consider a half line with origin in and directed along the unitary[4] versor $\hat{n} = \{n_x, n_y, n_z\}$. We impose that $\vec{P} + d \cdot \hat{n}$ lies on the surface of the cylinder, which in carthesian coordinates translates to:

$$d^2 n_x^2 \; + \; \left(d^2 n_y^2 \; - \; \frac{1}{2} \right)^2 \; = \; \frac{1}{4} \tag{1.22}$$

By solving equation 1.22 we can compute the length d of the portion of the half line contained in the cylinder as:

$$d \; = \; \frac{n_y}{n_x^2 + n_y^2} \tag{1.23}$$

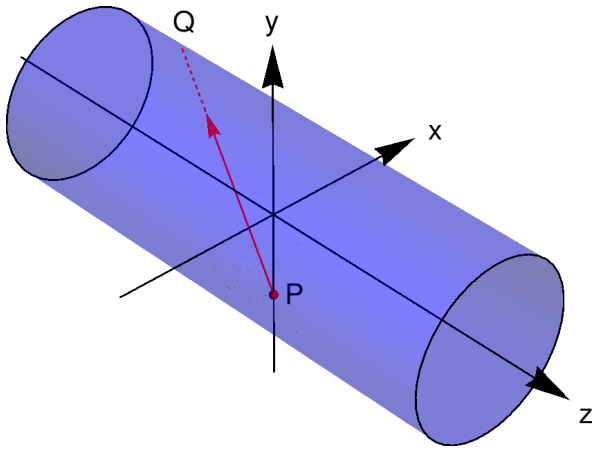


Figure 1.12: Geometric construction for the short theorem discussed in section 1.5.2

In polar coordinates:

[4] $|\hat{n}| = 1$

$$\begin{aligned} n_x &= \sin\theta\cos\phi \\ n_y &= \cos\theta \\ n_z &= \sin\theta\sin\phi \end{aligned} \tag{1.24}$$

substituting in 1.23:

$$d = \frac{\cos\theta}{\sin^2\theta\cos^2\phi + \cos^2\theta} \tag{1.25}$$

if we fix the value of θ and vary solely the value of ϕ the mean value of d can be written as:

$$\begin{aligned} \langle d\rangle &= \frac{1}{2\pi}\int_0^{2\pi} \frac{\cos\theta\, d\phi}{\sin^2\theta\cos^2\phi + \cos^2\theta} = \\ &= \frac{1}{2\pi}\Big[\arctan\big(\cos\theta\tan\phi\big)\Big]_{\phi=0}^{2\pi} = 1 \end{aligned} \tag{1.26}$$

Being the result independent of θ this can be drawn with any arbitrary distribution and the mean will consistently be 1.

QED ■

Using this property we can write the average distance d_B between successive hits (let's call it the *Bounce* distance) on the wall of a cylindrical channel of arbitrary radius r as $d_B = 2r$.

The value $2r$ works fairly well also as an approximation of the *Bounce* distance in a single dendritic channel; for the thermalisation in NCP the *Bounce* distance can be computed via Monte Carlo. Results of these simulations[5] are presented in table 1.2. The effect of using an entire NCP is to increase the *Bounce* distance due to the channels occasionally merging together. Varying the channel radii R_{Min} and R_{Max} while maintaining constant their density per unit of macroscopic surface shows that larger channels merge more frequently leading to further departure from the single dendrite case.

At any specific moment in the thermalisation process the average time between successive hits (*Bounce* time) can be computed as:

[5]Since dendritic channels have variable radiuses the distribution of hits on their walls is not uniform per unit of their length and has a distribution that is dependent on the local channel radius. Since also the *Bounce* distance depends on the local radius, to correctly compute the values in the table we employed the entire dynamic presented in section 1.5.3; we nonetheless find useful to introduce here these values.

Geometry	R_{Min}	R_{Max}	$\langle R \rangle$	d_B	$\frac{d_B}{\langle R \rangle}$
Cylinder	–	–	–	$2 \cdot \langle R \rangle$	2
Dendrite	3 nm	4 nm	3.5 nm	6.8 nm	1.94
Dendrite	4 nm	6 nm	5.0 nm	9.7 nm	1.94
Dendrite	5 nm	8 nm	6.5 nm	12.6 nm	1.94
Dendrite	7 nm	9 nm	8.0 nm	15.6 nm	1.95
NCP	3 nm	4 nm	3.5 nm	8.6 nm	2.46
NCP	4 nm	6 nm	5.0 nm	15.0 nm	3.00
NCP	5 nm	8 nm	6.5 nm	25.5 nm	3.92
NCP	7 nm	9 nm	8.0 nm	43.6 nm	5.45

Table 1.2: Computed *Bounce* distances for single dendrites and an NCP geometry generated using the standard parameters detailed in section 1.3.1. We propose also the ratio between the *Bounce* distance and the mean radius which in the case of a cylindrical channel should be exactly 2; the discrepancy of this parameter with the value 2 is a measure of how much the single channel can predict the $2\gamma/3\gamma$ annihilation ratio inside of channels.

$$t_B \;=\; \left\langle \frac{d_B}{v} \right\rangle \;=\; \langle d_B \rangle \cdot \left\langle \frac{1}{v} \right\rangle \;=\; 2r \left\langle \frac{1}{v} \right\rangle \tag{1.27}$$

where v is the positronium speed at the considered instant. Writing the mean of the product as the product of the means is allowed only since the distribution of the speeds and the *Bounce* distances are completely uncorrelated. We compute the factor $\langle 1/v \rangle$ from the simulations presented in section 1.5.1. Since this value depends on the number of interactions the positronium underwent, $t_B(n)$ will depend on the number of interactions n. We should note here that the average time Ps requires to undergo n interactions in the NCP can be obtained from the summation of the bounce times; since it will become useful later on, let's define:

$$t_F(n) \;=\; \sum_{i=0}^{n} t_B(i) \tag{1.28}$$

We can now describe which fraction of the positronium produced in the channels will not be decayed after it underwent n interactions and of the decayed fraction how much has decayed due to wall interactions into two γs and how much self-annihilated into three γs. We do so recursively by defining three series, $F_S(n)$, $F_{2\gamma}(n)$ and $F_{3\gamma}(n)$ that represent, after n interactions took place, the fraction of positronium that has survived, the fraction that has decayed into two γs and the fraction that has decayed into three γs. We impose:

$$F_S(0) \;=\; 1 \qquad F_{2\gamma}(0) \;=\; 0 \qquad F_{3\gamma}(0) \;=\; 0 \tag{1.29}$$

then, recursively[6]:

$$\begin{aligned} F_S\,(n+1) &= F_S(n)\cdot(1-p)\cdot e^{-t_B(n)/\tau} \\ F_{2\gamma}\,(n+1) &= F_{2\gamma}(n) \;+\; F_S(n)\cdot p \\ F_{3\gamma}\,(n+1) &= F_{3\gamma}(n) \;+\; F_S(n)\cdot(1-p)\cdot(1-e^{-t_B(n)/\tau}) \end{aligned} \tag{1.30}$$

where $\tau \approx 142\,\text{ns}$ is the lifetime of ortho-positronium in vacuum and p is, as previously defined, the annihilation probability during a single interaction with the wall. An example of the evolution of the three populations is shown in figure 1.13

We can use the series $F_S(n)$ to make a prediction of the positronium lifetime in capped channels. We can compute it as the average of the time $t_F(n)$ employed to perform n interactions (defined in equation 1.28) using the fraction of positronium that decays either to two or three γs between the nth interaction and the successive one to weight the average.

The fraction of positronium that decays between successive Ps-wall interactions is given by:

$$\begin{aligned} F_D(n) &= F_S(n) - F_S\,(n+1) \\ &= F_S(n)\left[1 \;-\; (1-p)e^{t_B/\tau}\right] \end{aligned} \tag{1.31}$$

We should note that:

$$\sum_{i=1}^{\infty} F_D(i) \;=\; 1 \tag{1.32}$$

[6]Be aware that since $t_B \ll \tau$ the numerically stable formulation of the recursive step of the 3γ fraction can be obtained through Taylor series expansion as: $F_{3\gamma}\,(n+1) = F_{3\gamma}(n) \;+\; F_S(n)\cdot(1-p)\cdot t_B(n)/\tau$; the surviving fraction should be computed by difference: $F_S\,(n+1) = 1 - F_{3\gamma}(n+1) - F_{2\gamma}(n+1)$, the definition of $F_{2\gamma}$ stays unchanged.

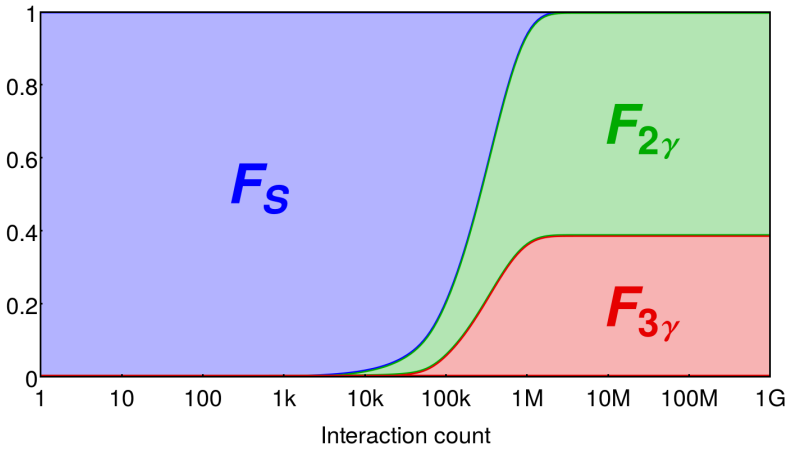


Figure 1.13: The surviving fraction of positronium (F_S), the wall-annihilating fraction ($F_{2\gamma}$) and the self-annhilated fraction ($F_{3\gamma}$) plotted on top of one another as a function of the number of interactions with the channel walls. The simulation has been performed assuming the wall to be at a temperature of 300 K, the channel radius to be 5 nm, the effective mass to be 28 amu and the annihilation probability during a single interaction to be $p = 1/600000$.

which is equivalent to say that in a capped channel all positronium atoms will eventually decay. We can now compute the average positronium lifetime in capped channels as:

$$\tau_C = \frac{\sum_{i=1}^{\infty} t_F(i) \cdot F_D(i)}{\sum_{i=1}^{\infty} F_D(i)} = \sum_{i=1}^{\infty} t_F(i) \cdot F_D(i) \tag{1.33}$$

The value of τ_C is dependent on the channel radius and on the *Bounce* time which is dependent on the thermalisation dynamic which is given, ultimately, by the sample temperature T and the effective mass M employed to simulate the exchange of energy between the channel wall and the positronium atom (as described in section 1.5.1).

Table 1.3 lists the expected positronium lifetime inside of capped channels for different values of the channel radius, M and T. Unpublished measurements of the positronium lifetime in NCPs with channel radius of $3 \div 4$ nm whose channels had been capped with TiO_2 showed a positronium lifetime τ_C in the order of 40 ns. If were to apply this model to such measurement we would expect $1/p$ to be in the order of 600000.

Saito and Hyodo [86] provide an estimate for the p value in the form of the quenching rate λ in a cavity. The work distinguishes between two

	T = 20K					T = 300K					
p = 1 / 150000	M (amu) \ r (nm)	3	5	8	12	20	M (amu) \ r (nm)	3	5	8	12	20
	16	**31**	**43**	56	67	79	**16**	11	18	26	**36**	50
	28	25	**35**	**46**	55	66	**28**	10	16	24	**32**	**45**
	40	21	29	**38**	**46**	56	**40**	9	14	21	29	**40**
	56	16	23	30	**37**	**46**	**56**	8	13	18	25	**35**
	80	12	17	22	28	**36**	**80**	7	10	15	21	30
	112	8	12	16	21	27	**112**	5	8	12	17	25
p = 1 / 300000	M (amu) \ r (nm)	3	5	8	12	20	M (amu) \ r (nm)	3	5	8	12	20
	16	55	71	84	94	104	**16**	23	**34**	**47**	60	76
	28	50	64	76	85	95	**28**	22	**32**	**44**	57	72
	40	**45**	58	69	78	87	**40**	20	30	**42**	53	68
	56	**39**	51	61	69	78	**56**	19	28	**39**	50	64
	80	**33**	**42**	51	59	68	**80**	17	25	**35**	**45**	58
	112	26	**34**	**42**	**49**	58	**112**	15	22	**31**	**39**	51
p = 1 / 600000	M (amu) \ r (nm)	3	5	8	12	20	M (amu) \ r (nm)	3	5	8	12	20
	16	82	97	108	115	121	**16**	**40**	56	72	85	100
	28	78	92	102	109	116	**28**	**39**	55	70	83	98
	40	74	87	97	104	111	**40**	**38**	53	68	81	95
	56	69	81	91	98	105	**56**	**37**	51	65	78	91
	80	62	74	83	90	97	**80**	**35**	**48**	62	74	87
	112	55	66	74	81	89	**112**	**32**	**45**	57	68	81

Table 1.3: Positronium lifetimes (in ns) in capped channels computed with different values for the channel radius, sample tempertaure and effective mass parameters. To increase readability we highlighted entries in the range 40 ± 10ns.

different processes that can induce the positron annihilation: wall pickoff and conversion of ortho-positronium into para-positronium due to paramagnetic centers present on the cavity walls. Working under the assumption of Ps trapped inside of cavities whose geometry provide a mean free distance $L = 500\,\text{nm}$, Saito and Hyoto[86] evaluate the former to have rate $\lambda_{wpo} = 6 \cdot 10^4 \text{s}^{-1}$ based on the data of[85] and the latter to have rate $\lambda_{wconv} = 2 \cdot 10^6 \text{s}^{-1}$ based on the data of[87].

If we convert these in their equivalent value for p assuming the positronium to be traveling (for consistency with [86]) at 10^5m/s we obtain:

$$p \;=\; \frac{v}{L \cdot (\lambda_{wpo} \,+\, \lambda_{wconv})} \;\approx\; \frac{1}{100000} \tag{1.34}$$

which is not consistent with our estimate. We ran both values through the computations illustrated in the sections that follow (and played around with the parameters) and found consistently that $1/p = 100000$ gives the best accordance with the experimental data. A possible explanation for the discrepancy comes from the fact that, as we will see later on from the permanence time, for most of the 40 ns that the positronium spends inside of the capped channels its temperature is close to that of the sample. It is expected [56] that quantum effects change the behavior with which low temperature positronium atoms interact with the channel walls; in all of the subsequent computations the effect of the annihilations against the channels walls are either relevant during phases in which the positronium has a high energy or they are small corrections to the result. We cannot, therefore, exclude that a lower value of p is needed to properly describe the behavior of cold positronium in nanoscopic channel than it is necessary with hot positronium.

1.5.3 Phase 3: Interaction count spectrum

We want to compute the interaction count spectrum $\mathcal{S}(n)$ defined in section 1.5 as the probability that an infinitely lived positronium will require exactly n wall interactions to exit the nanoscopic channel.

To compute $\mathcal{S}(n)$ we need to generate points on the nanoscopic surface of cylindrical channels, single dendrites and the NCP using the the effective implantation profile $\mathcal{P}_e(z)$. Since generating random points on complex surfaces using given PDFs is not trivial we'll resort, when necessary, to generating the points with whichever distribution is more handy and then to assign to each simulation the weight:

$$w \;=\; \frac{\mathcal{P}_e(z)}{\mathcal{P}_{Gen}(z)} \tag{1.35}$$

with z the depth of the generated positronium origin point, $\mathcal{P}_e(z)$ the effective implantation profile and $\mathcal{P}_{Gen}(z)$ the probability to generate a point at depth z using the two aforementioned methods. The distribution $\mathcal{P}_{Gen}$ can be numerically computed by means of sampling, binning and interpolation.

In both the aforementioned cases, after a point on the surface has been extracted, a random inward pointing direction is then drawn and a straight trajectory is traced from the Ps origin point up until the first surface hit; a new direction is then drawn and the process is repeated until the positronium exits the nanoscopic channel. The amount of interactions required to do so is then collected into a histogram that, filled using the weight stated in equation 1.35, constitutes $\mathcal{S}(n)$.

Interaction count profiles for entire NCPs generated using the parameters given in section 1.4 are shown in figure 1.14. To try to mimic the hit count spectrum obtained from the NCP geometry we attempted to use a single dendrite from the same NCP or a cylinder with a radius similar to that of the channel in the NCP; we saw that in both cases $\mathcal{S}(n)$ differs greatly from the NCP geometry and its proposed mocks. Instead we used table 1.2 to compute an *effective* radius for the NCP channels, i.e.: the radius that, employed to generate a cylindric channel or a single dendrite, would give the same *Bounce* distance. To give an example, for an NCP with radiuses comprised between 5 nm and 8 nm the effective radius is 12.8 nm. We show in figure 1.14 the hit profiles given by a cylindrical channel and a single dendrite generated with these parameters. The result is that the single dendrite is capable of closely mimicking the entire NCP, while the cylindrical channel fails to do so. We searched also a suitable effective radius for the cylindrical channel and found that an effective radius of 20 nm yields the result closest to the full NCP, but still comes short of mimicking it well enough to be a suitable substitute in simulations.

The bottom line is: we can employ a single dendrite with the correct effective radius to compute $\mathcal{S}(n)$ resulting, again, in a considerable reduction in computational complexity and time.

The NCP profile spectrum for an implantation energy of 7000 eV plotted in figure 1.14 shows a slight bump just above 1000 hits. This is due to the finite length of the channels causing part of the positronium traveling in the channels to be reflected back when it reaches the channel's end: when the reflected Ps atoms reach the surface of the macroscopic surface of the nanochanneled plate they raise the value of $\mathcal{S}(n)$ causing the bump.

Before computing thermalisation spectra for the positronium, there are two additional quantities that we can compute from the results obtained so far: the emission angle from the nanochanneled plate and the permanence time of positronium in the nanoscopic channels.

We define the emission angle of a positronium from the surface as the angle its trajectory is forming with the normal to the NCP's macroscopic surface

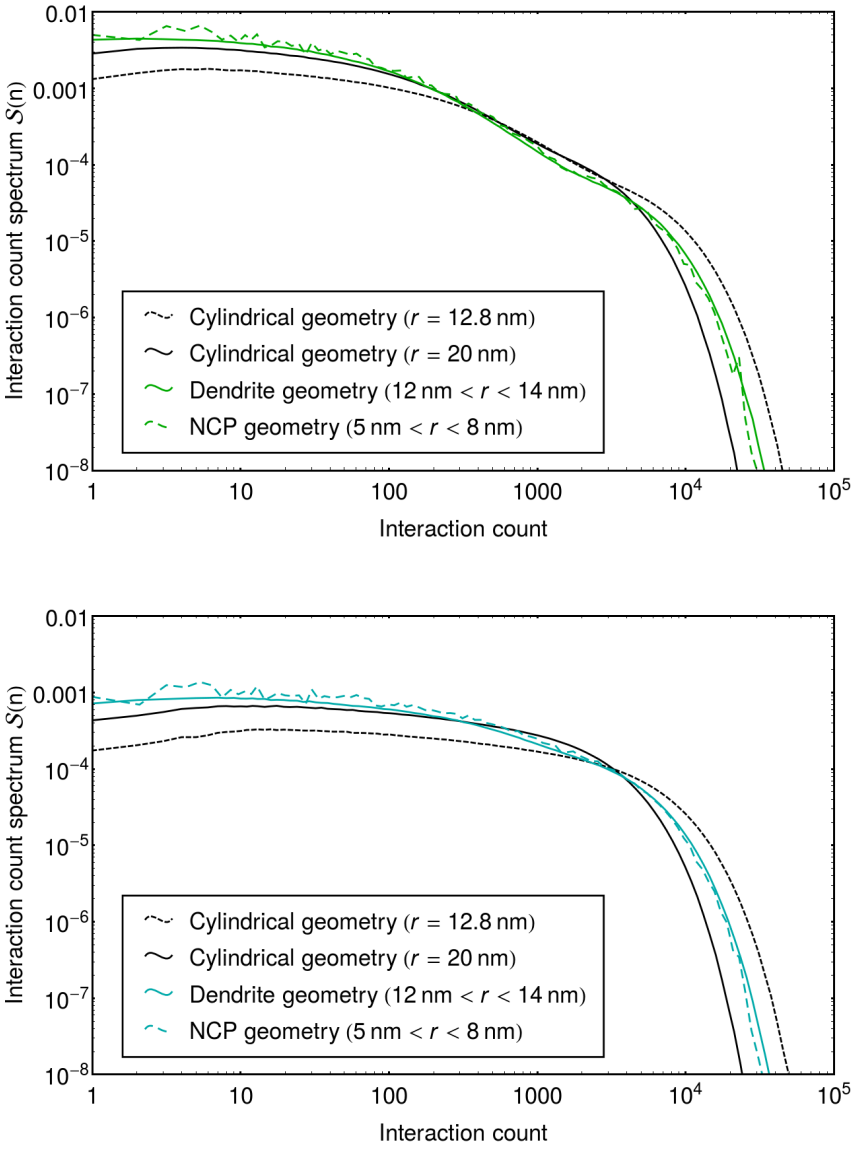


Figure 1.14: Interaction count spectrum within the NCP geometry (generated with the parameters given in section 1.4) for a 7 keV implantation energy (above) and 11 keV implantation energy (below). Overlaid in the same color is the spectrum given by single dendrites generated with radii that reproduce the *Bounce* distance measured in the NCP. Overlaid in black are the interaction count spectra of cylindrical channels with radius chosen, in one case, to mimic the NCP *Bounce* distance, in the other to reproduce as closely as possible the NCP spectrum.

the moment the positronium escapes into the vacuum. The distribution of the emission angles is computed as a byproduct of the Monte Carlo that we run to compute $\mathcal{S}(n)$ in the form of being the value of the randomly drawn propagation direction inside of the channels during the last iteration, i.e.: when the positronium reaches the channel's mouth. Since the emission angle distribution is only weakly correlated to the interaction count distribution it is irrelevant, as we verified numerically, whether particles emitted from the channels are weighted using their survival probability $F_S(n)$ (as should be if we aim to be rigorous) or not.

In chapter 8 of his PhD thesis[88], Dr. Caravita described how the experimental apparatus described in chapter 3 of the present work can be used to measure the angular distribution of the positronium emerging from an NCP target; we'll employ Dr. Caravita's results to validate (or invalidate) our model. In figure 1.15 the emission angle distributions for the three geometries are shown superimposed to Caravita's measurements[88].

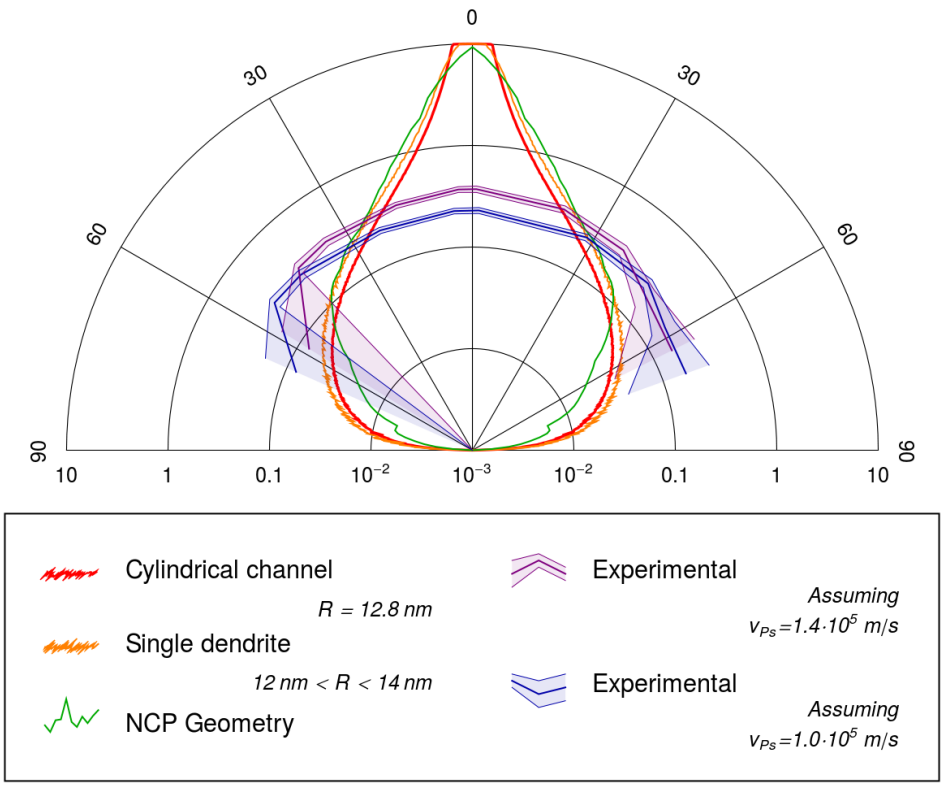


Figure 1.15: Positronium emission angle, measured from the macroscopic surface's normal. Experimental data displayed in Blue and Purple has been measured in the *Breadbox* apparatus (described in chapter 3) with the technique described in chapter 8 of [88].

It can be clearly seen from the plots that the distribution predicted by the model described insofar differs greatly from the experimentally measured one. This is the first of three experimental tests that this model is incapable of describing correctly. We'll illustrate the other two before proceeding in section 1.9 to explain how this model can be modified to correctly reproduce the experimental data.

1.6 Three gamma fraction

Another parameter that we are able to compute from the results obtained insofar is what portion $S_{3\gamma}$ of the implanted positrons will decay in the target or in its vicinity into three gamma rays instead of two. This parameter is useful to verify our model against experimental data since measurement of the emitted γ spectrum allows the distinction of the two decay modes and is easily done in laboratory [55, 48]. A measure of $S_{3\gamma}$ has been provided my Mariazzi et al. [55] and will be used here as a comparison for the values predicted by the model.

Let's start by considering solely the positronium that has been produced after a thermal positron has reached the wall of a nanoscopic channel. Some of that positronium will decay in 3γ either while flying in the channel or after it has escaped the NCP, some other will decay due to its interaction with the channel walls. The fraction of the positrons that decays into 3γ can be written as:

$$S_{3\gamma}^{\mathrm{Ps}} \;=\; \sum_{n=0}^{\infty} \mathcal{S}(n) \cdot [F_S(n) \;+\; F_{3\gamma}(n)] \tag{1.36}$$

as usual the sum can be approximated by truncation due to:

$$\lim_{N\to\infty} \sum_{n=N}^{\infty} \mathcal{S}(n) = 0 \tag{1.37}$$

On top of these 3γ decays we must add all of the Ps atoms produced by thermalized positrons that reach the macroscopic surface of the material (which become relevant at lower implantation energies).

To properly write the amount of Ps that is being produced (either in the channels or on the macroscopic surface) we need to account for the fact that not all of the thermalized positrons reaching the material surface are re-emitted as ortho-positronium. Let's call $\epsilon_{\text{o-Ps}}$ the fraction of the thermalized positrons that are re-emitted as ortho-positronium. We can write the fraction of 3γ-annihilating positronium as a fraction of the total implanted positrons as:

$$S_{3\gamma}^{\text{Tot}} \;=\; \epsilon_{\text{o-Ps}} \cdot \left(S_{3\gamma}^{\text{Ps}} \cdot f_{\text{Nano}} + f_{\text{Macro}} \right) \tag{1.38}$$

where f_{Nano} is the fraction of implanted positrons that reach the nanoscopic surface after thermalisation and f_{Macro} the fraction reaching the macroscopic surface after thermalisation (see figure 1.9 for reference). We are assuming here that the positrons that are not emitted as positronium will eventually decay into 2γ. This is certain for a fraction equal to $\epsilon_{\text{o-Ps}}/4$ (which is re-emitted as para-positronium) and a safe bet for the rest of the population.

Experimentally the 3γ decays have been determined as a fraction of the non backscattered, non epithermal positrons since these escape the geometric acceptance of the apparatus. We need, therefore, to normalize the fraction in equation 1.38 to account for that. We obtain finally:

$$S_{3\gamma} \;=\; \frac{\epsilon_{\text{o-Ps}}}{1 - f_{\text{Back}}} \cdot \left(S_{3\gamma}^{\text{Ps}} \cdot f_{\text{Nano}} + f_{\text{Macro}} \right) \tag{1.39}$$

where f_{Back} is the fraction of the implanted positrons that is either backscattered or returned epithermal (see again figure 1.9).

To compute $S_{3\gamma}$ we need to provide a value to $\epsilon_{\text{o-Ps}}$. Mills measure this parameter[48] as the limit of the ortho-positronium production efficiency from a silica surface as the implantation energy approaches zero obtaining values of $0.38^{+5\%}_{-10\%}$ and $0.44^{+5\%}_{-10\%}$ for differently-aged samples. The measure by Mariazzi et al. [55] provides a value of $0.42^{\pm 3.5}$ for $\epsilon_{\text{o-Ps}}$. Considering both measures we'll employ $\epsilon_{\text{o-Ps}} = 0.42$.

The computed value of $S_{3\gamma}$ for the cylindrical and single dendrite geometries is shown in figure 1.16. We do not need to employ the entire NCP geometry to attain these prediction since, as we observed in section 1.5.3, the single dendrite geometry can effectively mimic it at a considerably lower computational cost. A value of 1/100000 has been assumed for the pickoff probability p, a value of 28 amu for effective mass and a temperature of 300 K. It can be seen from the plot that the predicted value for the three gamma fraction does not reproduce the experimental data of[55]. The match can be bettered by modifying the simulation parameters, in particular by employing unrealistic values for p. We will resist this temptation for the moment, and differ all of these consideration to section 1.9.

1.7 Permanence time

Let's set aside the prediction failures shown in the last two sections and move on to expose another one. The positron implantation and thermalisation process takes place in a very short time scale (< 100 fs). On the contrary

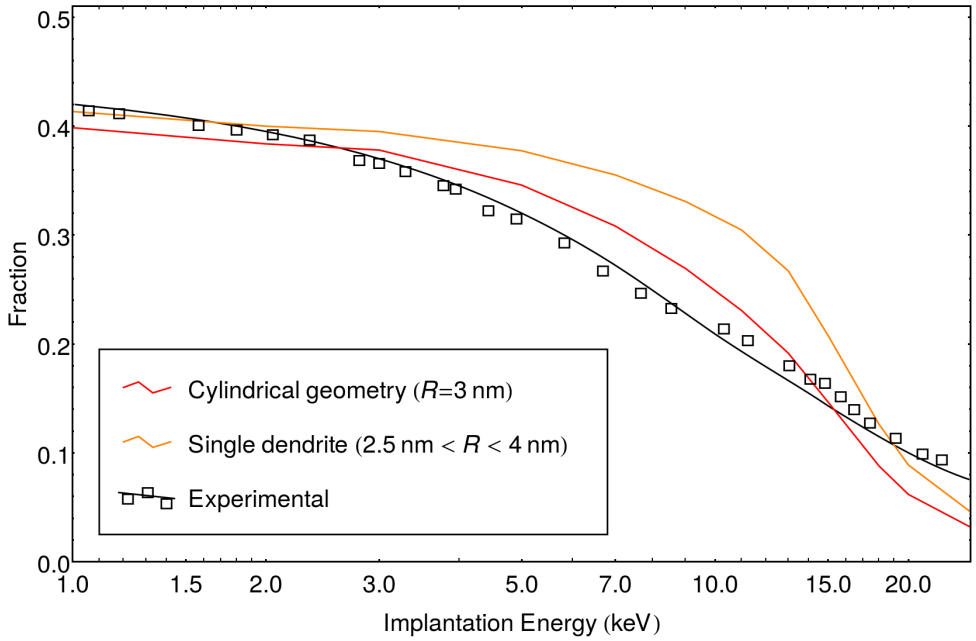


Figure 1.16: 3γ fraction as predicted by our model and as measured by Mariazzi et al. [55], the continuous black line is the fit to the experimental data presented in the original article. It is not necessary to test the full NCP geometry since as we say previously the single dendrite geometry mimicks it closely (and presents a widely different computational burden).

measurement of positronium production in pellets of compressed silica powder show thermalisation times in the order of $10 \div 20\,\text{ns}$ [89, 90]. Considering that the energy distribution of positronium emerging from nanochanneled plates displays a thermalized component, it would not be surprising if the permanence time of positronium inside of the nanoscopic channels were in the order of nanoseconds too.

We can compute the average permanence time of positronium inside of the nanoscopic channels from the average time $t_F(n)$ necessary to interact n times with the channel walls (see section 1.5.2), the fraction $\mathcal{S}(n)$ of positronium that takes n interactions to exit the channel (see section 1.5.3) and the fraction $F_S(n)$ of the positronium that is able to survive n interactions with the channel walls (see again section 1.5.2):

$$T_P \;=\; \frac{\sum_{n=1}^{\infty} \mathcal{S}(n) \cdot F_S(n) \cdot t_F(n)}{\sum_{n=1}^{\infty} \mathcal{S}(n) \cdot F_S(n)} \tag{1.40}$$

As usual infinite sums can be cut to large enough finite sums with negligible error being introduced in the process. The permanence time is extremely useful, as we will see, to properly interpret time-of-flight data, but unfortunately very little data is present in the literature to compare the predicted values to.

An experimental measure has been again provided by Mariazzi et al.[91]. Mariazzi distinguishes between thermalized and hot Ps atoms and proposes two different permanence times: $18 \pm 6\text{ns}$ for the thermalized positronium and $< 7\,\text{ns}$ for the hot positronium.

To properly compare our results with the paper we should be able to discriminate, based on the interaction count, which events should be classified as pertaining to the hot distribution and which to the thermalized ones. Measurements performed in similar conditions[55] claim the cold portion of the emitted positronium to sum to 19% of the total. The result of this phase is strongly dependent on this quantity; the article presents no estimation of the uncertainty of this value but presents it as a mere estimate; considering the whole procedure presented in[55] we will employ tentatively an uncertainty of $\pm 5\%$ to estimate the accuracy of our reconstruction. We will select from the population of positrons that exit the target the 19% that underwent the highest number of interactions and assume that it represents a suitable sample of the thermalized positronium (we'll call that *Cold* fraction). We can make this assumption due to the fact that the temperature of the positronium decreasing monotonically as the number of this increases; therefore the 19% Ps that underwent the most interactions is, necessarily also the coldest 19% independently from the sample temperature and of the value employed for the effective mass M. We will compute the mean permanence time over this fraction. We decided to perform the same computation also on the

second topmost 19% (*Middle* fraction) and on the remaining Ps population for reference *Hot* fraction.

The results are shown in table 1.4, with the thermalized fraction of[55] being determined to travel in the channel, assuming $1/p = 100000$, $1.76^{+1.2}_{-0.7}$ns with the error deriving from the experimental uncertainty over the amount of the cold fraction. Even assuming much higher values for $1/p$ cannot resolve the discrepancy. Once more we do not care about the apparent discrepancy and will defer the discussion to the end of this chapter.

1.8 Energy spectra

We show in figure 1.17 an example of spectrum computed using formula 1.19. The resulting curve, as is the case with all of the simulated spectra we analyzed, can be fitted surprisingly well with a combination of two thermal distributions, one having an energy of several thousand degrees and another having temperature close to room temperature. This is in excellent accordance with the previous literature that several times found the emitted positronium distribution to consist, apparently, of the sum of two distinct thermal populations [55, 91, 92]. We will attempt to provide an explanation of the phenomenon in section 1.10; for the scope of this section we will restrict ourselves to characterize phenomenologically the behavior.

We expect the kinetic energy distribution of particles in a thermalized gas to follow the probability distribution:

$$\mathcal{D}(E,T) \;=\; \frac{2}{\sqrt{\pi}}\left(\frac{1}{k_B T}\right)^{3/2} \cdot \sqrt{E}\; e^{-E/(k_B T)} \tag{1.41}$$

with k_B being Boltzmann's constant, T the temperature of the gas and E the kinetic energy of the molecule. We employ equation 1.41 to formulate a proper fit model for the simulated spectra as:

$$\mathcal{M}(E) \;=\; f \cdot \mathcal{D}(E,T_1) \;+\; (1-f)\mathcal{D}(E,T_2) \tag{1.42}$$

We fitted the spectra obtained from the simulator employing the model in 1.42 with the constraints: $0 \leq f \leq 1$ and $T_1 < T_2$. The result of the fit process for a single spectrum is shown in figure 1.17. Due to the computed values presenting no experimental uncertainty we cannot provide a χ^2 to evaluate the fit; instead we'll provide the L^2 norm of the function difference normalized to the L^2 norm of the simulated spectrum:

$$L^2_{Norm} \;=\; \frac{\int_0^{0.2\,\mathrm{eV}} \left(\mathcal{M}(E) - \mathcal{K}(E)\right)^2 dE}{\int_0^{0.2\,\mathrm{eV}} \mathcal{K}(E)^2 dE} \;=\; 0.0012 \tag{1.43}$$

Implantation energy	$\frac{1}{p}$	Total	Hot	Middle	Cold
7 keV	$10 \cdot 10^3$	29.3 ps	6.2 ps	23.9 ps	92.4 ps
7 keV	$20 \cdot 10^3$	61.2 ps	8.3 ps	34.7 ps	0.21 ns
7 keV	$50 \cdot 10^3$	0.19 ns	11.6 ps	56.8 ps	0.72 ns
7 keV	$100 \cdot 10^3$	0.36 ns	14.4 ps	80.6 ps	1.43 ns
7 keV	$200 \cdot 10^3$	0.56 ns	16.9 ps	0.11 ns	2.21 ns
7 keV	$300 \cdot 10^3$	0.65 ns	18.1 ps	0.12 ns	2.61 ns
7 keV	$500 \cdot 10^3$	0.75 ns	19.3 ps	0.14 ns	3.0 ns
7 keV	$600 \cdot 10^3$	0.78 ns	19.6 ps	0.14 ns	3.11 ns
7 keV	10^6	0.84 ns	20.3 ps	0.15 ns	3.35 ns
7 keV	∞	0.94 ns	21.5 ps	0.17 ns	3.75 ns
11 keV	$10 \cdot 10^3$	66.3 ps	16.8 ps	60.1 ps	0.2 ns
11 keV	$20 \cdot 10^3$	0.16 ns	28.7 ps	0.11 ns	0.52 ns
11 keV	$50 \cdot 10^3$	0.49 ns	55.6 ps	0.27 ns	1.73 ns
11 keV	$100 \cdot 10^3$	0.88 ns	84.5 ps	0.47 ns	3.19 ns
11 keV	$200 \cdot 10^3$	1.28 ns	0.11 ns	0.71 ns	4.62 ns
11 keV	$300 \cdot 10^3$	1.47 ns	0.13 ns	0.83 ns	5.3 ns
11 keV	$500 \cdot 10^3$	1.65 ns	0.15 ns	0.97 ns	5.94 ns
11 keV	$600 \cdot 10^3$	1.7 ns	0.15 ns	1.01 ns	6.12 ns
11 keV	10^6	1.81 ns	0.16 ns	1.09 ns	6.5 ns
11 keV	∞	1.99 ns	0.18 ns	1.25 ns	7.13 ns

Table 1.4: Permanence time in the NCP. A single dendrite geometry has been employed with radius $2\,\text{nm} < R < 4\,\text{nm}$, a sample temperature of 300 K and an effective mass of 28 amu. Temperature and effective mass have relatively little influence over the permanence time. The permanence time has been calculated as a mean over all of the emitted positronium (*Total*) as the mean of highest 19% portion of the interaction count distribution (*High*), as the mean of the second topmost 19% (*Middle*) or as the mean of the lower 62% (*Low*). We can see that for higher implantation energies (11 keV the permanence time of the *Cold* fraction almost doubles, this is due to more deeply implanted positronium requiring more interactions to exit the channel.

where $\mathcal{K}(E)$ is the simulated energy spectrum and $\mathcal{M}(E)$ its fit.

The trending of the fit parameters for a variety of effective masses and channel radii is shown in figure 1.18. It can be noted that the model allows for the production of a thermal fraction of positronium with a temperature close to that of the target, nonetheless this requires either the employment of extremely low values for the effective mass ($M < 1\,\text{amu}$) or extremely thin channels. Experimental evidence[93] shows that a significant portion of cold positronium can be obtained even with channels with radii as large as 9 nm.

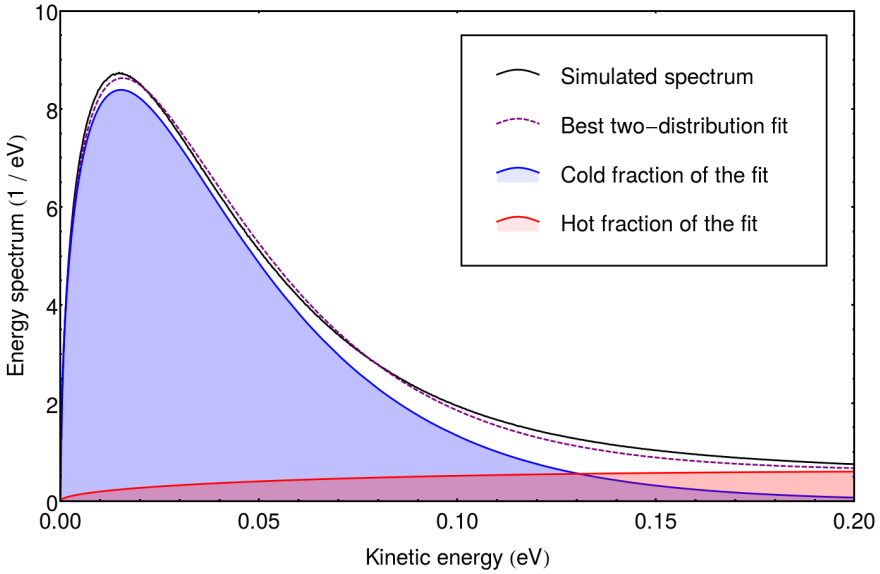


Figure 1.17: Simulated Ps kinetic energy spectrum (in black), overlaid in dashed purple the model of equation 1.42 fitted upon the simulated spectrum. The two components of the model employed to fit the spectrum have also been plotted in red and blue. The simulation has been performed using dendritic channels with $4\,\text{nm} < R < 6\,\text{nm}$, an effective mass of 4 amu a sample temperature of 300 K, implantation energy 7 keV and a wall annihilation probability of 1/100000. The low value for the effective mass has been chosen to provide an energy distribution closer to full thermalisation, the rest of the parameters for computational convenience. From the fit 52.8% of the simulated Ps atoms belong to the cold component (354 K) and the rest to the hot one (6003 K).

We would like to make a note here. Our code allows us to compute the final energy spectrum over a window of 1 eV which is large enough to store the totality of the spectrum features; the fit procedures here described have been run over the entire window. If instead we restrict the window to a smaller

region (such as $0 \div 0.2\,\mathrm{eV}$) and perform again the fit procedure we observe that the temperature of the hot fraction will decrease greatly by factors that range from 5 to 15. This impacts the analysis of the experimental data of most ToF experiments since the experimental data usually spans only a limited window; the aforementioned effect will cause them to underestimate the temperature of the hot fraction.

1.9 A different scattering model

In section 1.5 we formulated a model for the thermalisation of positronium inside of the nanoscopic channels. In the following sections we found the model to repeatedly fail to correctly reproduce experimental data.

If we closely analyze the various discrepancies we see that they all point towards a single direction: the positronium appears to escape the nanoscopic channels after undergoing too few interactions. Namely: the low value for the effective mass M found in section 1.8 is due to the fact that Ps atoms would need to lose more energy during each collision than what would be expected from Sauder's model. The low permanence time shows that the particle is emitted early from the channels; lowering the temperature of Ps atoms inside of the channels by lowering M and lowering p to zero cannot account for this discrepancy: more interactions are required to match the experimental data. Finally the dependency of the $2\gamma/3\gamma$ fraction on the implantation energy shows little increase of the 2γ component with the increased implantation depth up until the end of the nanoscopic channels is hit, meaning that the increase in the hit count against the walls does not cause a significant increase in the annihilations against the channel walls. To compensate for this the parameter $1/p$ should be lowered to unrealistic values; even if we fix the macroscopic trend this way, the shape of the measured and predicted curve will differ greatly.

If we try to figure out how the interaction count might be wrong we could decide to doubt the validity of the geometric model: in fact employing smaller channels will increase greatly the average of the interaction count spectrum. This solution is liable to bring the model closer to the experimental data, but is incapable to address the wrong emission angle spectrum.

In section 1.5 we decided to employ a uniform distribution to draw the direction of positronium atoms leaving the channel wall after an interaction. This was due to the channel wall being knowingly irregular and to us lacking data regarding the exact shape of the wall: it was a reasonable and convenient model but, it appears now, to have also been wrong. We will replace now the angular distribution of Ps atoms leaving the wall with a parametric distribution; we chose the distribution so that it is easily computationally synthesized and so that the parameters allows us to make it more forwardly

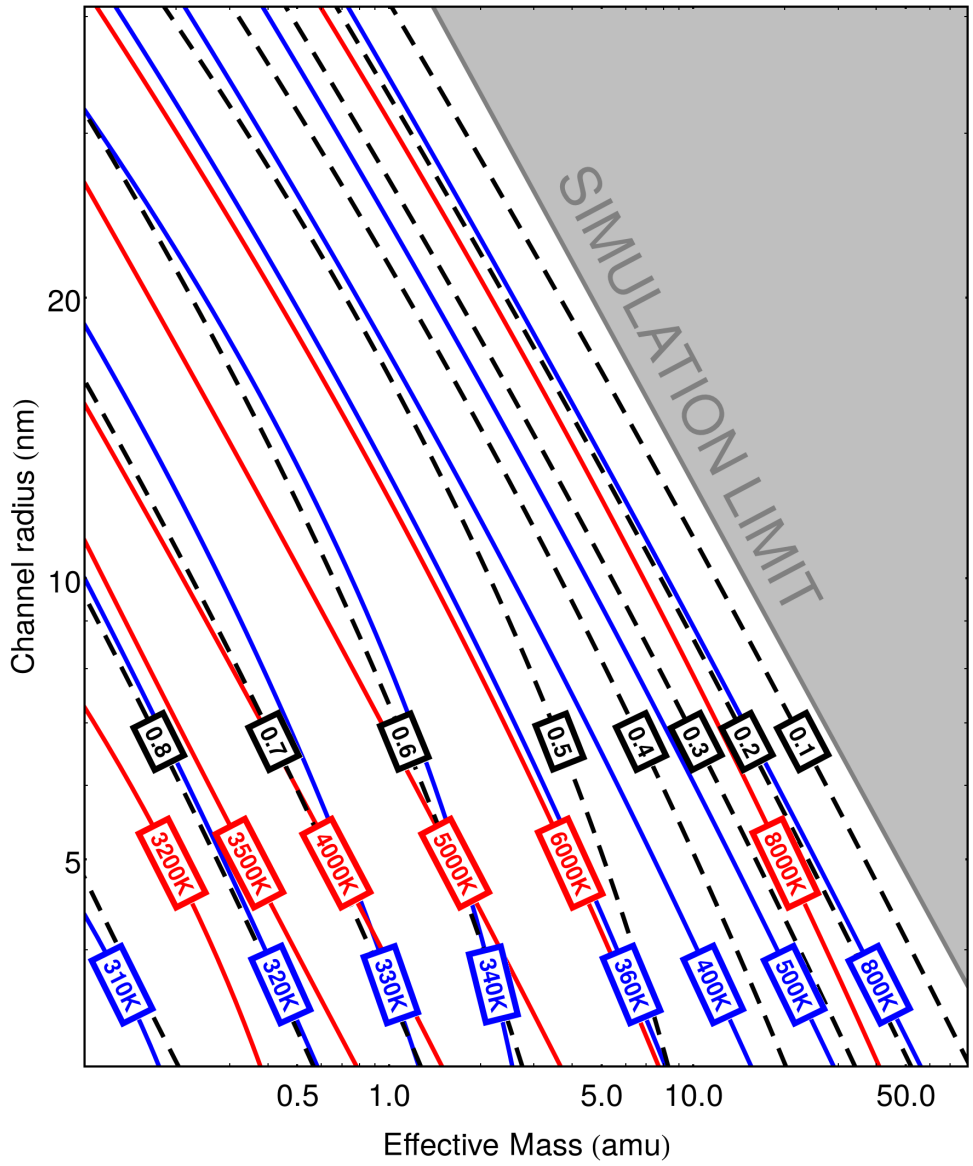


Figure 1.18: Fit parameters for spectra simulated with implantation energy 7 keV, sample temperature of 300 K, $1/p = 100000$ and dendritic channels whose mean radius is given by the vertical axis and radius spread is $R_{Max} - R_{Min} = 2$ nm. The fit model is presented in equation 1.42. In red the temperature T_2 of the hot component, in blue the temperature T_1 of the cold component and in black the cold positronium fraction f. To allow for the production of a significant fraction of cold positronium a very thin radius needs to be employed, which is not consistent with experimental evidence [93].

or laterally distributed.

We employ a polar coordinate system centered on the channel wall and whose axis is orthogonal to the channel wall itself. We then draw the polar angle ϕ from a uniform distribution and generate the azimuthal angle θ by computing:

$$\theta \;=\; \cos^{-1}\left(R^{\,1/q}\right) \tag{1.44}$$

where R is a random variable uniformly distributed in the interval $[0, 1]$ and q is the parameter we employ to decide how forwardly distributed the distribution should be. For $q = 1$ this distribution is uniform over the 2π solid angle. For $q > 1$ the distribution becomes forward-heavy and for $0 < q < 1$ the distribution becomes lateral-heavy. In figure 1.19 we show the shape of the distribution for different values of q.

Due to the distribution in ϕ being uniform, theorem 1 from section 1.5.2 holds its validity; the bounce distance in cylindrical channels will therefore be independent of the value of q; the same is not guaranteed for NCPs and single dendrites. We present in table 1.5 the recomputed the *Bounce* distances d_B in the new angular distribution model for NCPs and single dendrites with $q = 3$. We notice from the table that, to all practical purposes, values of d_B computed with $q = 3$ are indistinguishable from values of d_B computed with $q = 1$.

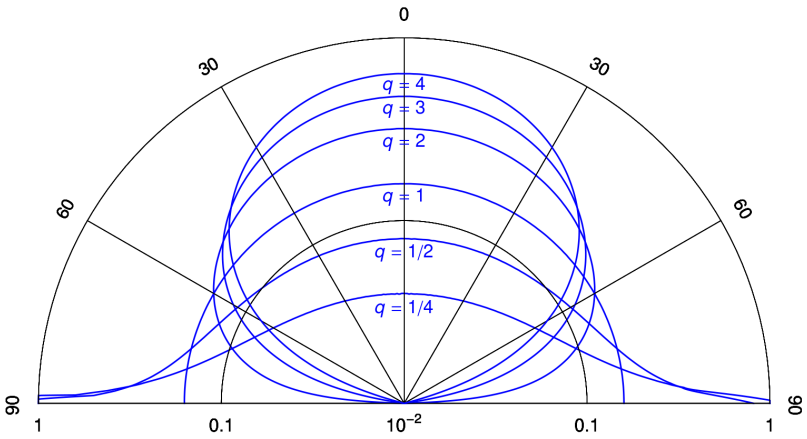


Figure 1.19: New angular distribution model for the positronium atoms bouncing off the wall of a nanoscopic channel for different values of q.

By trying different values of q we found that $q = 3$ fairly reproduces the experimental data. Let's review shortly the results presented in the previous chapters. First of all we will compare the interaction count spectra computed

with $q = 1$ (uniform distribution) and $q = 3$ (the new angular distribution). Figure 1.20 shows comparison of the old spectra with the recomputed ones. As can be clearly seen, the number of collisions has increased considerably while, at the same time, most features of the spectra are conserved (including the bump due to the reflection on the channel's bottom end). Once again the single dendrite geometry employed with an effective radius is capable of properly mimicking the interaction count spectrum given by the full NCP geometry.

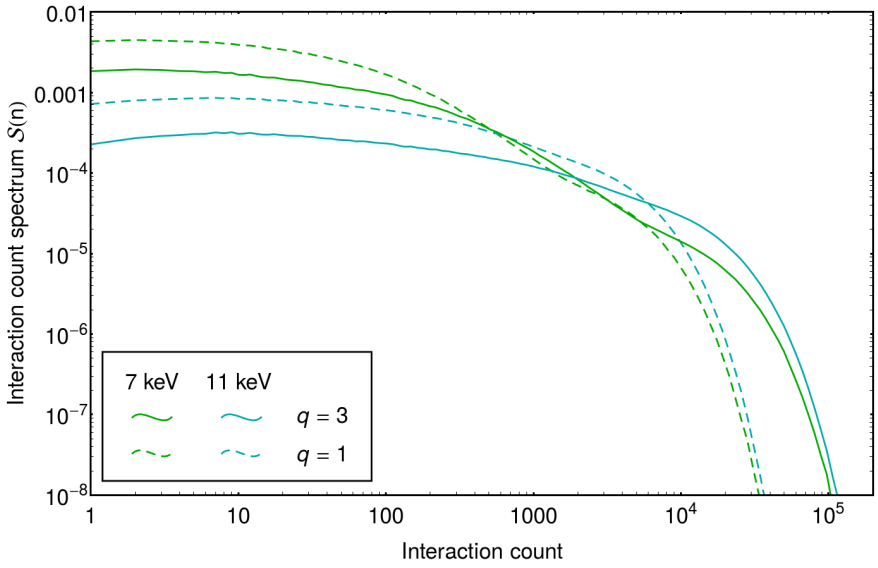


Figure 1.20: $\mathcal{S}(n)$ profiles for the single dendrite geometry ($12\,\text{nm} < R < 14\,\text{nm}$) for $q = 1$ (uniform distribution) and $q = 3$ (new angular distribution).

Secondly we compute the angular distribution for the new angular distribution model with $q = 3$. Figure 1.21 shows the new angular distribution which, in particular when employing the full NCP geometry, presents a much better agreement with the experimental values.

We proceed, then, to compute the 3γ fraction with the new angular distribution (as usual with $q = 3$). The result (shown in figure 1.22) shows, without the need to tune its parameters, an extremely good agreement with the experimental data up to an implantation energy of 9 keV. By shaving off a fraction of a nanometer from the channels radii (or, alternatively by setting $1/p = 78000$ and the production efficiency at the surface to 0.41) it is possible to make the prediction given by the dendritic geometry to pass within most of the square markers of the plot in figure 1.22, but we find more compelling the agreement given without tuning.

Geometry	R_{Min}	R_{Max}	$\langle R \rangle$	d_B	$\frac{d_B}{\langle R \rangle}$
Cylinder	–	–	–	$2 \cdot \langle R \rangle$	2
Dendrite	3 nm	4 nm	3.5 nm	7.0 nm	1.94
Dendrite	4 nm	6 nm	5.0 nm	10.0 nm	1.94
Dendrite	5 nm	8 nm	6.5 nm	13.3 nm	1.94
Dendrite	7 nm	9 nm	8.0 nm	16.1 nm	1.95
NCP	3 nm	4 nm	3.5 nm	8.7 nm	2.48
NCP	4 nm	6 nm	5.0 nm	14.8 nm	2.96
NCP	5 nm	8 nm	6.5 nm	25.7 nm	3.95
NCP	7 nm	9 nm	8.0 nm	46.0 nm	5.75

Table 1.5: Computed *Bounce* distances for single dendrites and an NCP geometry computed similarly to those in table 1.2 but using the new angular distribution with $q = 3$.

Despite our best attempts to reproduce it by tweaking the simulation parameters, it appears that this model cannot justify the presence of the bump displayed by the experimental points above the 10 keV implantation energy. Considering that this change of curvature in the experimental data is located in a region where the portion of positronium that reaches thermal energies becomes relevant[54], it is possible that the excess in 3γ annihilations is due to the interaction between the positronium and the channel walls changing its behavior as a consequence of quantum effects not being negligible anymore, as predicted in[56, 55].

We computed again, this time by using the new angular distribution model with $q = 3$, the permanence times of positronium inside of the nanoscopic channels (see table 1.6) similarly to what we did in section 1.7. As we did in section 1.7 we divided the Ps atoms population in three components based on the interaction count and computed the permanence time separately for all the three components. The new value of $3.9^{+1.2}_{-0.8}$ns for the permanence time in conditions similar to that of [55] lies at 2.3σ from the experimental value, which is better than the value found in section 1.7 which lied at 2.7σ but still off the experimental value. Still apparent from the table is that an increase in the value of $1/p$ can provide better-fitting values, in particular $1/p = 300000$ brings the permanence time within one sigma of the experimental value. Notice

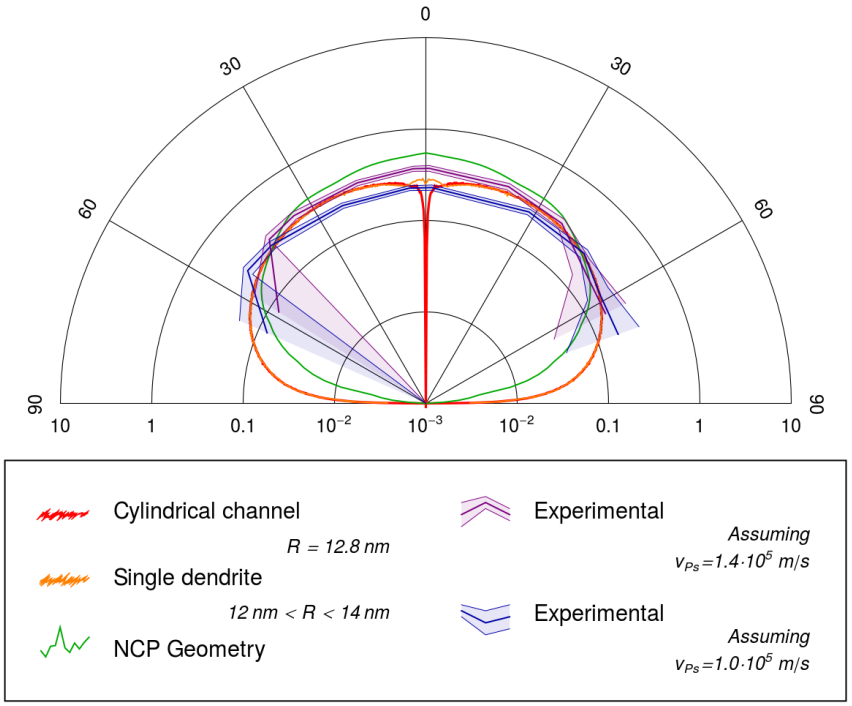


Figure 1.21: Positronium emission angle, measured from the macroscopic surface's normal as in figure 1.15 here computed using the new angular distribution with $q = 3$.

that the same conclusion could not be drawn from the results in table 1.4 since even a value of $p = 0$ cannot yield realistic permanence times. This apparent need for a higher $1/p$ adds up to the evidence that cold positronium seems to experience a lower value of p than the hot one (see again the lifetime value of positronium in capped channels computed in section 1.5.2, computation which is independent of q).

Finally we provide, in figure 1.23, the fit parameters landscape (as previously shown in figure 1.18) with the new angular distribution model. Firstly it can be noted that in the revised model the thermal component of the energy distribution approaches the sample's temperature for higher values of the effective mass M, placing it closer to the $16 \div 112$ amu interval that, as previously stated, we deem reasonable for the parameter M. The most notable difference with respect to the previous model is, nonetheless, the fact that the cooling capability of the NCP is held for higher values of the channel radius. Let's consider, as an example, the region around $\langle R \rangle = 5$ nm and $M = 28$ amu: an uniform scattering angle in the channel yields a cold component with a temperature of 800 K consisting in about 10% of the total, while the corrected model with $q = 3$ yields a cold component with a temperature

Implantation energy	$\frac{1}{p}$	Total	Hot	Middle	Cold
7 keV	$10 \cdot 10^3$	48.9 ps	12.2 ps	45.1 ps	0.15 ns
7 keV	$20 \cdot 10^3$	0.1 ns	18.6 ps	73.1 ps	0.34 ns
7 keV	$50 \cdot 10^3$	0.34 ns	30.0 ps	0.14 ns	1.25 ns
7 keV	$100 \cdot 10^3$	0.82 ns	41.2 ps	0.22 ns	3.18 ns
7 keV	$200 \cdot 10^3$	1.74 ns	55.1 ps	0.35 ns	6.9 ns
7 keV	$300 \cdot 10^3$	2.48 ns	64.5 ps	0.47 ns	9.89 ns
7 keV	$500 \cdot 10^3$	3.52 ns	76.9 ps	0.68 ns	14.06 ns
7 keV	$600 \cdot 10^3$	3.89 ns	81.3 ps	0.77 ns	15.54 ns
7 keV	10^6	4.85 ns	92.7 ps	1.05 ns	19.34 ns
7 keV	∞	7.23 ns	0.12 ns	2.05 ns	28.6 ns
11 keV	$10 \cdot 10^3$	83.0 ps	21.7 ps	76.1 ps	0.25 ns
11 keV	$20 \cdot 10^3$	0.22 ns	40.6 ps	0.16 ns	0.74 ns
11 keV	$50 \cdot 10^3$	0.94 ns	96.5 ps	0.48 ns	3.4 ns
11 keV	$100 \cdot 10^3$	2.37 ns	0.19 ns	1.36 ns	8.6 ns
11 keV	$200 \cdot 10^3$	4.81 ns	0.43 ns	3.85 ns	16.68 ns
11 keV	$300 \cdot 10^3$	6.57 ns	0.67 ns	5.95 ns	22.17 ns
11 keV	$500 \cdot 10^3$	8.87 ns	1.08 ns	8.69 ns	29.1 ns
11 keV	$600 \cdot 10^3$	9.65 ns	1.23 ns	9.61 ns	31.42 ns
11 keV	10^6	11.61 ns	1.65 ns	11.92 ns	37.16 ns
11 keV	∞	16.17 ns	2.71 ns	17.22 ns	50.33 ns

Table 1.6: Permanence times proposed in table 1.4 have been recomputed after applying the alternative positronium-wall interaction model with $q = 3$.

of 450 K consisting in 30% of the total, which is more realistic but still slightly hotter than expected.

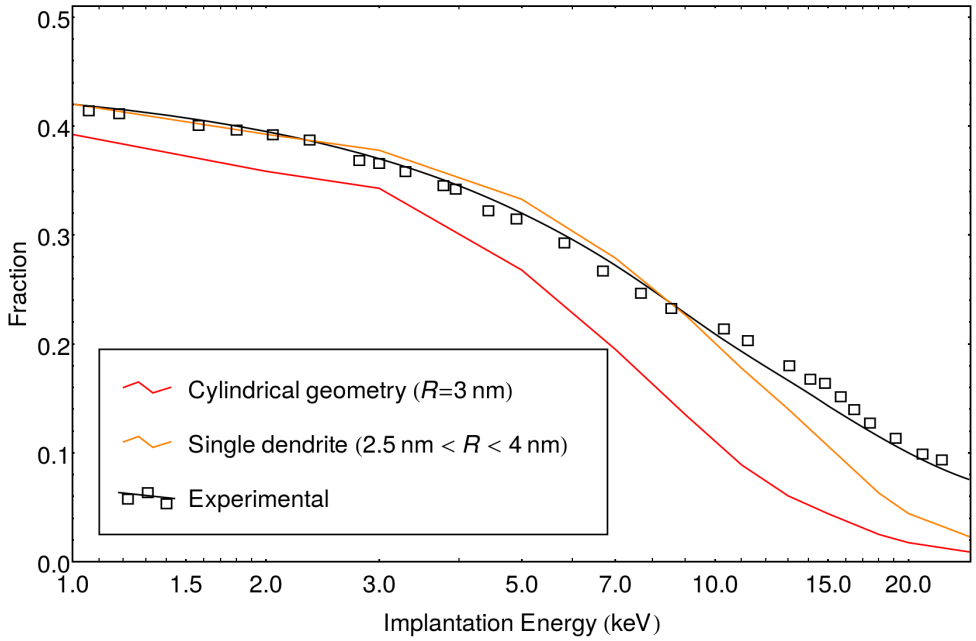


Figure 1.22: 3γ fraction as predicted by our model with the corrected distribution and as measured by Mariazzi et al. [55], similarly to figure 1.16.

To understand this result we need to consider the fact that there are two competing processes at work in the simulator. The $q = 3$ parameter decreases the average movement along the channel direction that each particle undergoes between successive interactions, thus increasing the number of particles that require lots of interactions to exit the channel (see figure 1.20). At the same time the annihilation against the walls suppresses the high-interaction-count component of the emitted positronium spectrum. In the case of the landscape depicted in figure 1.18 the limiting factor to the positronium thermalisation is the number of interactions required to escape the channel. Instead for the lowest values of $\langle R \rangle$ (< 5 nm) in the corrected landscape (figure 1.23) the limiting factor is the suppression of the coldest Ps atoms by annihilation against the channel walls, which is the mechanism underlying the change in trend happening in the bottom portion of the plot.

Colder and more abundant cold fractions can be obtained in the corrected model by lowering the value of p. As an example, lowering the value of p to $1/p = 300000$ will cause the aforementioned region to predict a low temperature component of 390 K and comprising 46% of the total. Alternatively, again

with $1/p = 300000$, a channel with $2\,\text{nm} < R < 4\,\text{nm}$ will yield, simulating with $M = 28\,\text{amu}$ a cold component of 360 K comprising once again 46% of the total.

The bottom line of these considerations is that the corrected model is capable of reproducing the experimental data much more closely than the non corrected one. What the corrected model is still incapable of explaining is the presence of the change of slope in the high-energy region of 3γ fraction (figure 1.22) and in general seems to need a lower value of p than that present in literature to best predict the temperature of the cold fraction of the Ps atoms. Both these deviation from the predicted behavior point toward the presence of a dependency of the parameter p on the Ps atom temperature. We have not enough data to strongly claim the validity of any model that employs a temperature-dependent value of p, still we will show in section 1.11 that at least one such model exists capable of explaining all of the available experimental data.

1.10 Explaining a two-population spectrum

Previous measurements of the energy spectrum of positronium emitted by NCPs[55, 91, 92] observed the energy distribution to consist of two thermal distributions with widely different temperatures: a colder one close to the temperature of the NCP sample and another ranging in the order of thousands of degrees. To our knowledge no explanation has yet been provided for the presence of two different thermalized populations.

Spectra simulated with the model illustrated in this chapter support this observation, being themselves excellently fitted by the combination of two thermal distributions; we will attempt to provide an explanation of the presence of two thermal distributions by deconstructing the thermalisation of positronium in the channels as simulated by our model. To provide a clear explanation we will need to refer often to illustrated material; we provide all the necessary plots in figure 1.24.

Firstly let's consider a simulation in which we implant positrons with an energy of 7 keV in an NCP with channels with effective channel radius $R = 10\,\text{nm}$ held at a temperature of 300 K. We'll let the positronium scatter inside of NCP channels with an effective mass $M = 28\,\text{amu}$ through interactions whose angular distribution is generated with $q = 3$. We'll set the wall annihilation factor to $1/p = 100000$. The choice of these parameters has been operated considering ease of computation[7], and as far as we can tell is irrelevant in terms of the final conclusions.

We start by dividing the produced positronium into three families depending on the number of interaction the positronium underwent before leaving

[7] And aesthetic outcome, of course.

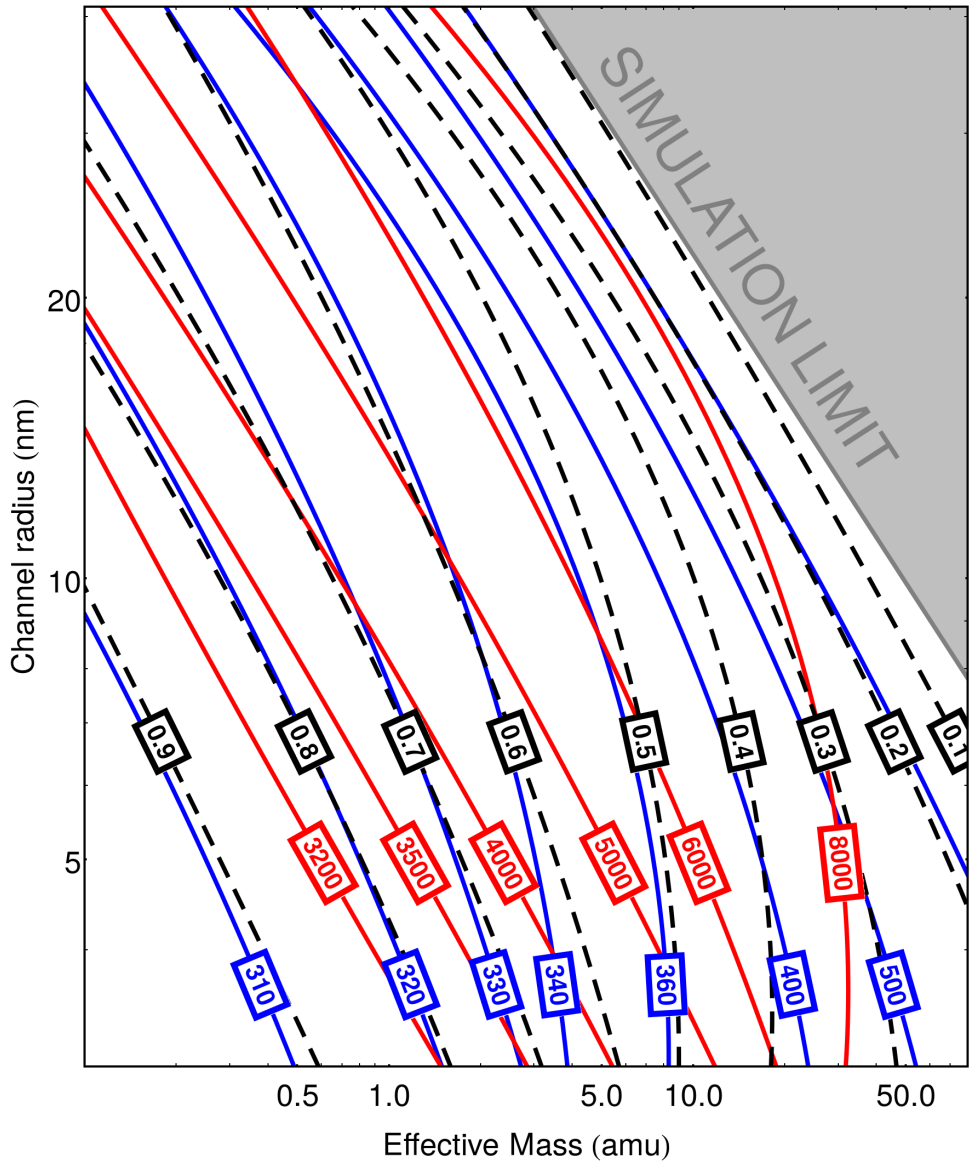


Figure 1.23: The fit parameter landscape presented in figure 1.23 recomputed with the corrected model.

the channel. Instead of employing sharp cuts to divide the population we will employ two transition zones where the population is divided between two different families. Panel a) of figure 1.24 shows the proportions with which we divided the population between a portion that undergoes few interactions before escaping the NCP (*Hot* distribution, in red), a portion that undergoes a large number of interactions before escaping (*Cold* distribution in blue) and a middle-ground portion (*Mid* distribution in purple).

Panel b) of figure 1.24 shows how the interaction count spectrum $\mathcal{S}(n)$ gets divided among the three portions. Since we decided to partition the positronium that escapes the channels instead of the positronium that has been produced inside of the channels, stacking the three distributions shown in the panel amounts not to $\mathcal{S}(n)$, but a portion is missing: it is the positronium that decays before escaping to the vacuum.

Panel c) shows the evolution of a distribution of Ps atoms that underwent the same amount of interactions after they were produced inside of the nanoscopic channel. As the thermalisation progresses the production distribution (the δ located at 3 eV) spreads out turning in a larger distribution resembling (as consequence of the central limit theorem) the normal distribution. This distribution then slowly travels toward the thermal distribution and, approaching it, becomes again more compressed. As the equilibrium distribution is more closely approached, the distribution loses its symmetry settling into its final shape. We colored the successive phases of the thermalisation according to the distribution (*Mid*, *Cold* or *Hot*) to which we assigned most the Ps atoms that have reached such thermalisation step when they left the NCP.

Finally in panel d) we see how the three portion contribute to the final shape of the positronium spectrum. The *Cold* portion consists mainly of Ps atoms that have completely reached thermalisation and, therefore, sums up to a distribution that closely resembles a thermal distribution with a temperature similar to that of the NCP sample, although the contribution of incompletely thermalized spectra makes it appear slightly higher than that of the sample. The *Hot* portion sums to a heavy tail located completely at high energies and contributes nothing to the regions that are usually measured experimentally (the dashed vertical line marks 0.2 eV which is a fair indication of the region most ToF experiments are sensible in).

The *Mid* portion consists of spectra that just reached the energy range that is visible by the instrumentation. Due to the thermalisation process slowing down in proximity of the zero of the energy scale and of the spectra becoming more compressed, the *Mid* portion sums to a shape that is heavier at lower energies and, therefore, mimics a thermal distribution. Being in the $0 \div 0.5$ eV region not influenced by the *Hot* portion, data analysis in this region will observe the spectrum to be composed of two thermal distributions. We depicted in dashed blue and purple the best fit to the *Cold* and *Mid*

contributions with thermal distributions with temperatures 600 K and 6600 K respectively.

1.11 Temperature-dependent wall annihilations

As anticipated several time across the chapter, we have three major indications that the parameter p, the wall annihilation probability during a single interaction, cannot be properly described by a constant: they are the shape of the 3γ fraction, the permanence time and the Ps lifetime in capped channels. As said before this amount of data does not allow us to determine what the correct dependence of p on the positronium energy, still one question remains open: does a model exist describing the dependency of p on the positronium temperature capable of reproducing the experimental data?

In an attempt to build such model we will employ one of the simplest dependency models available: a sharp step. We will set a threshold for the positronium energy and simulate the wall interactions employing one among two different values of p, depending on whether the energy of the positronium energy lies above or below the energy threshold. We need to find suitable values for the threshold and the two values of p. To estimate the high-energy p we can employ the Ps lifetime in capped channels. As stated in section 1.5.2 to explain a lifetime of 40 ns in capped channels assuming realistic effective mass values $1/p$ should have a value around ≈ 600000. Since for most of the time it spends inside of capped channels positronium holds the same temperature as that of the sample, we can assume $1/p = 600000$ as a good estimate for the value of p at low energy.

To obtain the other two parameters (p at high energy and the threshold) we can fit the experimental 3γ fraction profile shown in figure 1.25. The result is a value of $1/p = 28000$ for high temperatures and a threshold set at 0.40 eV. In figure 1.25 we show, superimposed to the experimental 3γ fraction, the 3γ fraction simulated with a fixed $1/p = 100000$ and the 3γ fraction simulated with a variable $1/p$: the fit shows a much better accordance with the presence of a bump at higher implantation energies.

If we proceed to compute the lifetime in capped channels using the newly-minted variable p model we obtain 44 ns, while for the permanence time the new model yields $17.6^{+4.5}_{-3.2}$ns for the cold fraction. Both values are in beautifully good agreement with the experimental data. This does not mean that the variable p model is correct on physics standpoint but only that the experimental data can be fitted[8] by this kind of model. As we will see at the end of chapter 2, this will be useful to know when trying to interpret ulterior experimental data.

[8] As a matter of fact, probably over-fitted

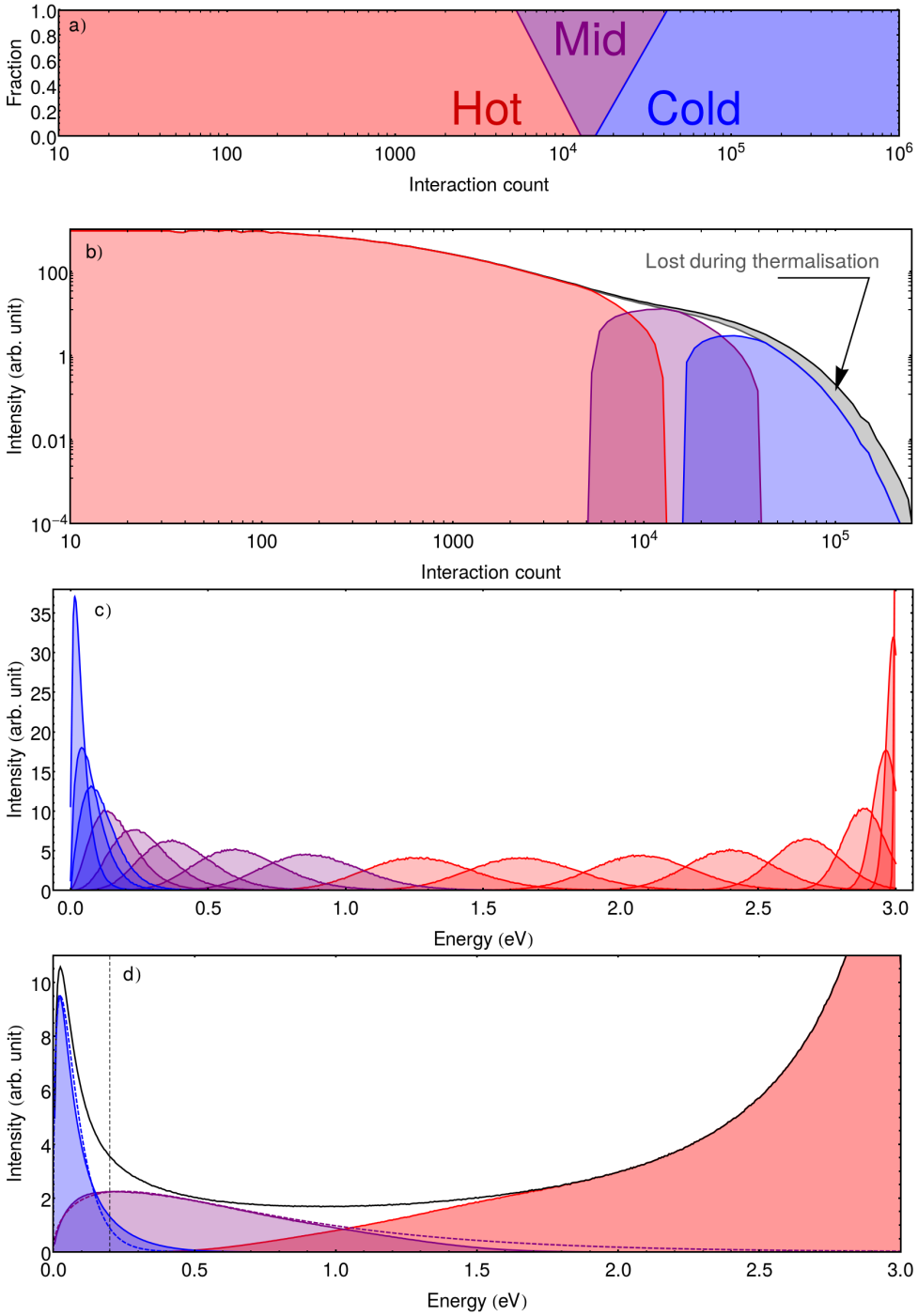


Figure 1.24: Illustration of the mechanism that leads to the positronium spectrum to show two thermal distribution. Please refer to section 1.10 for a detailed description.

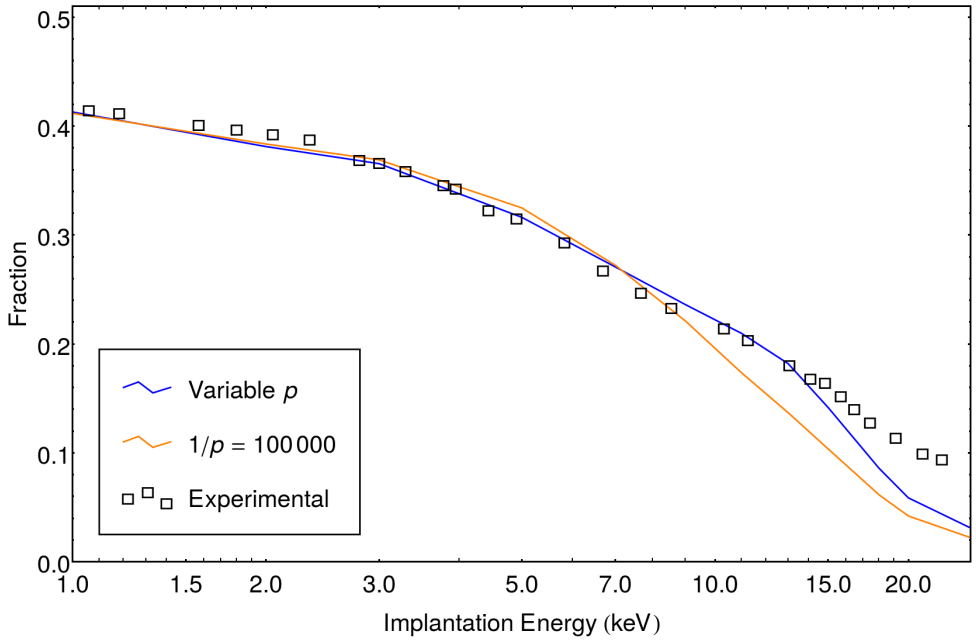
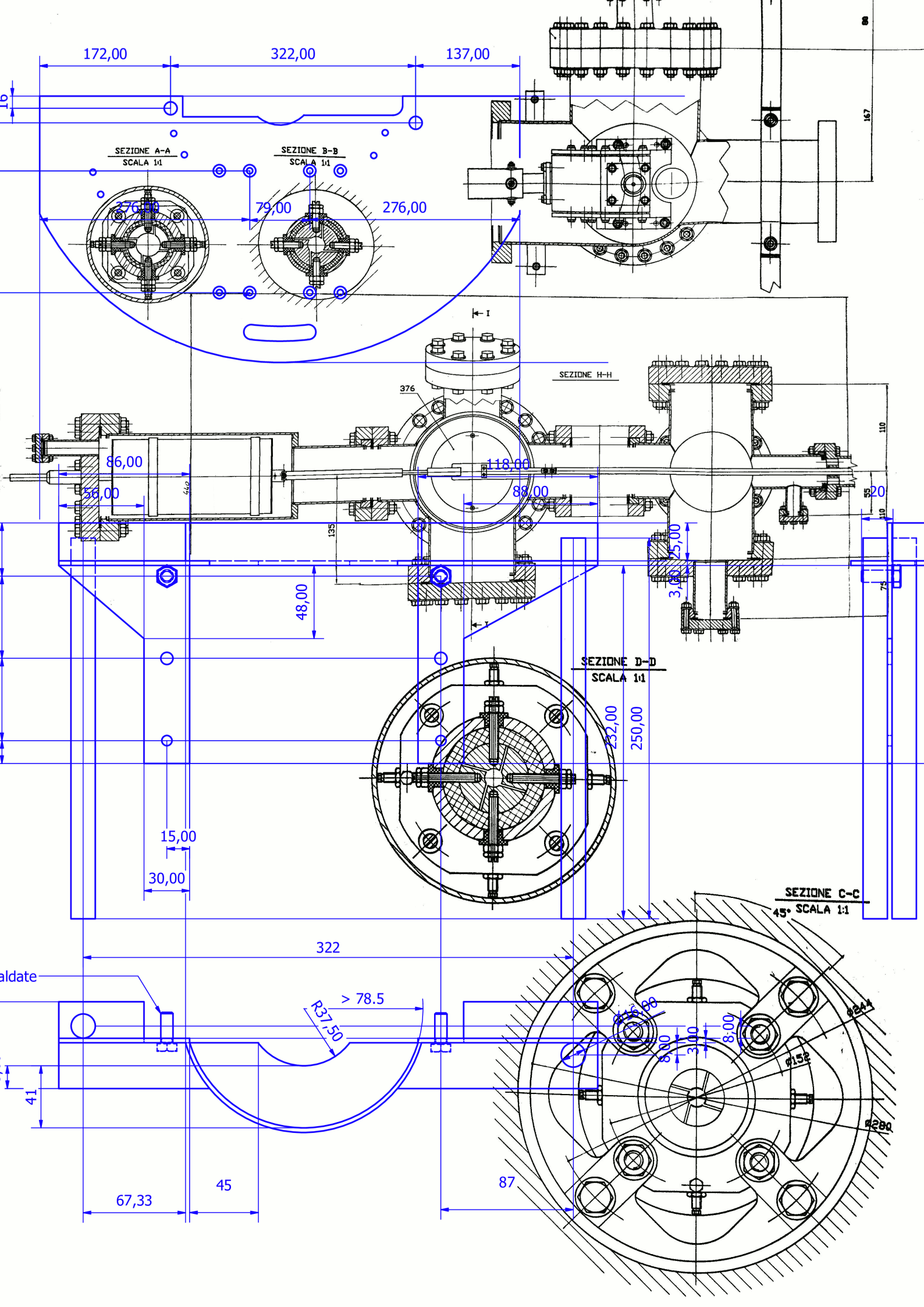


Figure 1.25: The experimental 3γ fraction with, superimposed, the fraction forecast by the constant p model and by the variable p model, showing much better accordance with the variable p model. While the limitation in the 3γ fraction at high energies for the fixed p model is mainly due to the positronium annihilating on the channel walls, in the variable p model it is caused by the limited length of the channels. The flex in the simulated curve at 15 keV is justifiable with the adoption of an insufficient channel length.

In black: details of the mechanical drawings from the original commissioning of the SURF experiment [38]; in blue detail from the mechanical drawings of the upgrade of the apparatus to PRINCEPS.


172,00
322,00
137,00
SEZIONE A-A
SCALA 1:1
SEZIONE B-B
SCALA 1:1
276,00
79,00
276,00
SEZIONE H-H
376
86,00
56,00
118,00
88,00
48,00
SEZIONE D-D
SCALA 1:1
232,00
250,00
15,00
30,00
SEZIONE C-C
45° SCALA 1:1
322
> 78.5
R37,50
41
67,33
45
87
ø244
ø152
ø200

Chapter 2

Time of Flight from Nanochanneled Plates

This was an environment built, not for man, but for man's absence.

– J.G. Ballard
High-Rise

We spent the last chapter to describe a method to simulate the positron conversion that takes place inside of nanochanneled plates. Being able to take a look at the phenomena that take place inside of an NCP is interesting by itself and extremely useful when attempting to develop new positron to positronium converters; still if no verification of the simulation results is performed against experimental data the whole effort would remain nothing more than an elaborated abstract speculation. We'll see here how we were able to characterize the velocity of positronium atoms produced through a nanochanneled plates. The nature of our measurement is similar to the one realized in 2010 by Mariazzi et al [55] on NCPs and to the one realized in the same year by Crivelli et al. [94] on thin mesoporous silica films.

2.1 The PRINCEPS experiment

In 1996 a continuous positron beam was commissioned at the positron laboratories in Trento[38, 95] to constitute the experiment called SURF. The SURF beam originates from a radioactive ^{22}Na source whose spontaneous β^+ decay produces high energy positrons with an energy spectrum that ranges from zero to 546 keV; the high-energy positron beam is then directed against

a 1 µm thick tungsten foil in which part of it is slowed down to thermal energy and re-emitted by from the foil surface with an energy of a few electronvolts. The beam is then accelerated with an electrostatic potential of 200 V and the resulting beam, consisting of high energy positrons that weren't slowed down by the tungsten and thermalized positrons accelerated to 200 eV are then directed onto an electrostatic deflector that operates a speed selection. After it has passed the deflector the beam, that began with a horizontal orientation, is now directed downward and consists only of its 200 eV component. A series of electrostatic lenses forms then the beam into the final spot onto the target sample to be studied. An electrostatic potential of several kilovolts, maintained between the entire beam apparatus and the target, determines the implantation energy with which the positrons are implanted in the target. This electrostatic potential is generated keeping the target referenced to ground and rising the potential of the rest of the apparatus to the desired voltage, this characteristic of the SURF apparatus constitutes an innovation that grants more freedom in the choice of target, allowing the use of mobile targets attached to manipulators or cryostats and heaters to control the target's temperature. A schematic representation of the SURF beam is shown in figure 2.1.

The SURF beam presents a spot with a FWHM of less than 1.5 mm and a pencil angle of less than 2^o. The overall efficiency of the apparatus has been measured[38] as $\epsilon \approx 2.7 \cdot 10^{-5}$, which means that we expect a beam intensity of $1000 e^+/s$ for a 1 mCi ^{22}Na source.

The original SURF experiment was designed to perform surface studies on samples held at cryogenic temperatures employing channeltrons to detect positrons and secondary electrons and Germanium detectors to measure γ ray energies and with this setup it has been employed for almost 20 years[96, 97]. We aim to employ the same beam to measure the energy spectrum of positronium produced from a nanochanneled target, to do so we upgraded the system by placing a new ^{22}Na source with an intensity of 45 mCi and by installing a new experimental chamber and a renewed detector and data acquisition system.

We called the renewed experimental apparatus PRINCEPS (Passage and Reflection in Nanochannelplates, Convertible Experimental Positronium Source). The name references the fact that the upgraded system was designed to allow both the measurement of Ps energy spectra from targets that produce positronium in reflection and target that produce Ps in transmission, category that is gaining ever increasing interest lately[98].

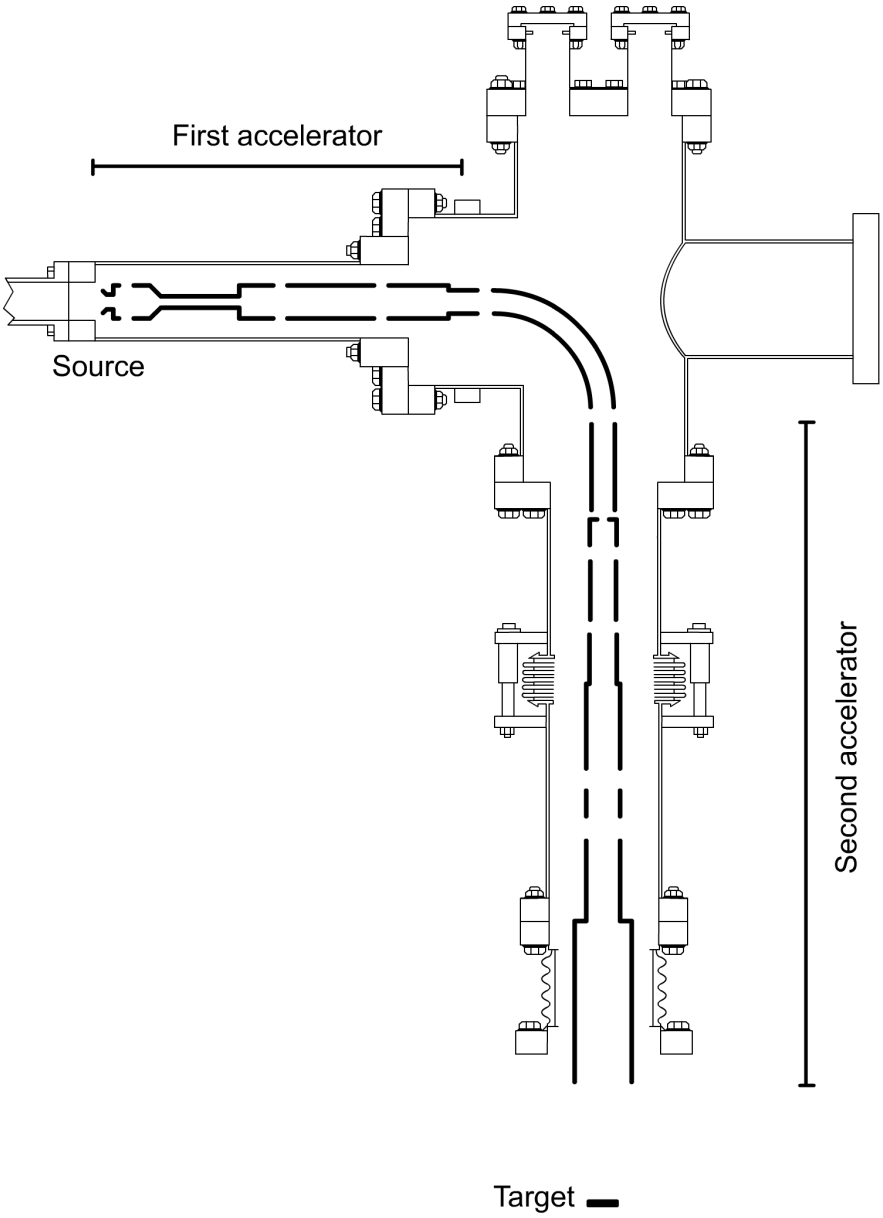


Figure 2.1: Sectional view of the SURF apparatus displaying its bent beam line connecting the source to the sample. The experimental chamber is not shown.

2.2 Experimental chamber geometry

The main design of the PRINCEPS experiment in fairly simple. A beam of positrons is implanted onto a target kept at a controlled temperature. We know that every implanted positron will cause, on average, 0.7 electrons from the target to be emitted when the positron collides with the Silicon Oxide surface [99, 100]; these electrons are collected by two micro channel plates (MCP) detectors facing the target[1] and covering a solid angle of $\approx 0.85\,\mathrm{sr}$. Some of the implanted positrons will then be re-emitted from the sample as positronium and decay in flight inside of the chamber. Employing NaI(Tl) scintillator detectors and a lead shielding shaped to form a slit we selectively detect Ps decaying in a volume whose distance from the target is as close to a fixed value as conceded by geometric contraints and statistic necessity. We chose five cylindrical 25.4×25.4mm PMT-coupled NaI(Tl) crystals as our detectors of choice to collect γs generated from Ps annihilations. Whenever an implanted positron causes a secondary electron to be expelled from the surface and be collected by an MCP and the positronium deriving from the positron conversion decays in the selected window and one of its emitted γs is collected by an NaI(Tl) detector we can estimate the time of flight (ToF) of Ps from the moment it is emitted from the sample to the moment it has decayed as the time elapsed between the detection of the electron and the detection of the γ ray. This is allowed since we expect both from thermodynamics and previous measurements[101], the positronium to have a speed around 10^5m/s, if we place the shielding to allow detection of annihilations in excess of 5 mm of distance from the target the positronium ToF will be in the order of magnitude of 50 ns. The distance between the target and the MCPs is in the order of 1 cm, the secondary electron is expected to have an energy in the order of a few tens of electronvolts [102] so we expect the time of flight of the secondary electron to be in the order of magnitude of a few tenths of a nanosecond. Similarly the stopping, conversion of the positron is a process that takes place in the time range of hundreds of femtoseconds. The time required taken by the positronium thermalisation in the NCP channels can be estimated to be in the range of a few nanoseconds (see 1.7). The distance between the annihilation area and the NaI(Tl) detectors ranges below 15 cm so the time of flight of the annihilation gamma adds up 0.5 ns to the total uncertainty. It is therefore reasonable to assume the time elapsed between the secondary electron and the gamma detection as a good approximation of the positronium time of flight but for the permanence time in the nanoscopic channels which is the only systematic that we'll have to take into account.

To realize the experimental chamber in which the sample is held and the positron conversion takes place we modified a chamber that we had built for a previous experiment. The chamber consists mainly in a cylindrical shaft with

[1] Photonis 34767

a radius of 75 mm capped on the two sides with CF100 flanges and realized with a 2 mm thick non magnetic steel. The central portion of the cylinder, for a total height of 140 mm has been reduced on one side of the cylinder to a radius of 30 mm with planar steel surfaces welded to join the two cylindrical portions. On the opposite side a complex of four flanges of different size permits the installation of the sealed passthroughs that connect the detectors and instrumentation installed inside of the chamber.

To adapt this chamber to the operation of PRINCEPS we performed two modifications. First of all we added two metal plates on the sides of the planar joints between the cylindrical sections of different radius; these additions will allow us to install the holder plane that houses the NaI(Tl) detectors and their necessary lead shielding. We opted to install these structures to experimental chamber in lieu of directly the support structures for the shielding and detectors to allow for an easier future repurposing of the experimental chamber.

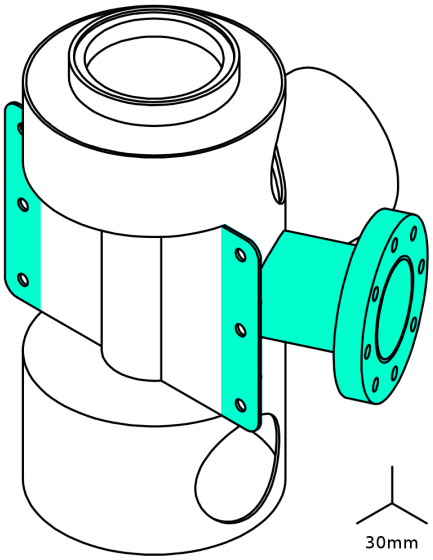


Figure 2.2: An isometric rendition of the experimental chamber employed in the PRINCEPS experiment, with the experiment-specific addition highlighted in aquamarine.

We then added a fifth CF63 flange on the cylinder side, centered on the vertical position in which the positron beam is expected to focus when the chamber is installed. This flange allows the horizontal mount of a cryogenerator that culminates in a highly thermally conductive sample holder that can be used to transmission targets. This is an unforeseen mounting point for the cryostat (the original design demanded it to be mounted on the bottom CF100 flange). Mounting the cryostat on the side leaves the region of space

below the sample free which allows for measurement of Ps production in the transmission scheme. At the same time the lateral cryostat mounting creates some additional difficulties to the alignment of the sample with the beam. It is therefore optimal to mount the cryostat on the bottom flange when producing Ps in the reflection scheme and on the lateral flange when employing the transmission scheme: the modifications to the chamber allow measurement in both setups, hence the name PRINCEPS (standing for Passage / Reflection In Nanochannels, Convertible Experimental Positronium Source). The chamber modifications are shown in figure 2.2.

Then we realized the supporting structure that holds in place the NaI(Tl) detectors and the lead shielding. This structure needs to answer to several performance requests: it has to be structurally sound enough to hold up to 100 kg of lead shielding, it has to be adjustable in its height with a range of a few centimeters with a precision of a fraction of a millimeter and in its pitch and roll angles to about one milliradian; finally since the alignment is performed without the shielding installed the structure needs to be rigid enough to avoid bending out of calibration when the lead shielding is mounted onto it. The solution we found to achieve this was to build a steel frame that bolts onto the experimental chamber that holds with the shortest possible arm two 16 mm hardened steel metrically filleted bars. An aluminum plate, 10 mm in thickness, repurposed from a previous experiment is concatenated to the steel bars via through holes and vinculated to the bars by bolts. Adjusting the bolts position by a fraction of a turn allows the calibration of the plate height and of its roll angle to the required precision while, at the same time, the elastic modulus $< 2 \cdot 10^{11}$Pa of the steel bars grants that the deflection under the applied load will be less than 10 µm which is compatible with the requirements. To hold the rear of the plate and adjust the pitch angle we opted for an adjustable angled prop whose length is also made adjustable by means of a metric 18 mm bolt. There is no need for the prop to be rigid enough to keep the pitch angle unchanged while the lead is being loaded since the calibration can be performed afterwards, still we expect the pitch deflection to be under 100 µrad. A CAD rendering of the structure is shown in figure 2.3.

2.3 Data acquisition and control

As said before the main parameter that we need to measure is the ToF of positronium that is produced on the target and annihilates in front of the shielding slit. The apparatus electronics provide the necessary functions to gather a start signal from the MCP detectors, a stop signal from the NaI(Tl) detectors and, whenever these arrive within a few microsecond window, generates a pulse signal whose amplitude is proportional to the difference

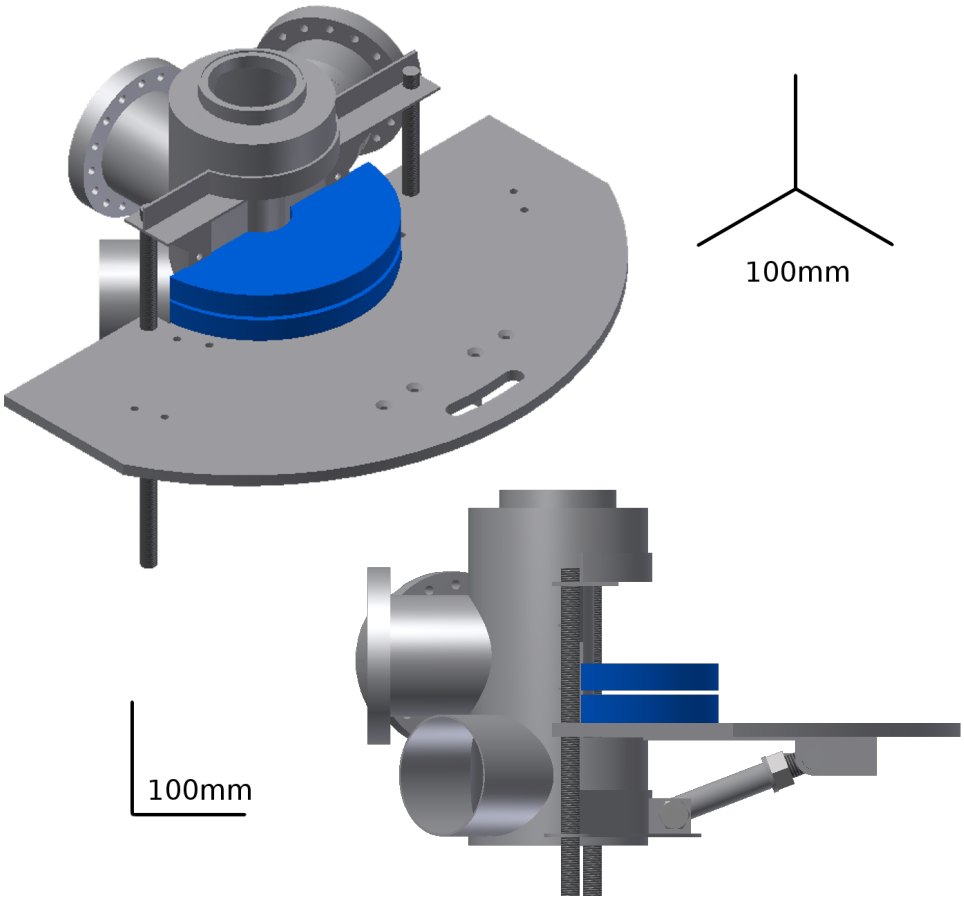


Figure 2.3: An isometric rendition of the detector holder viewed both and a lateral view of the same. In blue the internal lead shielding that masks the acceptance angle of the NaI(Tl) detectors and selects only a portion of the chamber. In the complete arrangement the five NaI(Tl) detectors are arranged around the blue shielding and are covered with a pile of lead bricks to minimize their background.

in time of arrival of the two signals. The NaI(Tl) detector signal intensity provides information about the energy of the detected γ ray; it is customary use to form the ToF signal employing only γs detected within an energy window whose boundaries are compatible with the physics of the Ps three-γ decay. Since, instead, we have the chance to record digitally the intensity of the γ annihilation corresponding to each ToF pulse and apply a windowing later, we opted to apply the least windowing possible though at an hardware level to, instead, apply the correct window later on. This choice allows us to choose the best performing energy window after the measurement has been concluded avoiding, in this way, the risk of wasting an entire campaign of measurements in the case of an incorrect windowing. We will discuss this in detail in section 2.6, still it might be interesting to present now a typical energy spectrum coming from our NaI(Tl) detectors like we do in figure 2.4. The preamplified signals coming from all five NaI(Tl) detectors are added together and need to be acquired alongside the ToF signal; of these two signals we need to register both the pulse height and the time of arrival.

We should mention now that for diagnostic purposes we installed another detector on the apparatus. At target height, on the side of the chamber, we placed a 76.2×76.2mm cylindrical NaI scintillator shielded on the sides by a tungsten tube with 10 mm-thick walls. Task of this detector is to constantly monitor the intensity of the positron beam, to pursue this goal we need to acquire the rate of this signal within reasonably short windows (in the order of seconds). The same applies to the raw signals coming from the MCP detectors, whose rate compared to the beam intensity in needed to properly calibrate the bias voltage of the MCP detectors (and is an extremely useful diagnostic to detect catastrophic discharge within the MCP detectors that we employed to set up an interlock to protect our hardware).

To acquire the ToF and γ energies we employed a commercial MCA (Multi Channel Analyzer), whereas to acquire signal rates we employed a custom built multiscaler, the design, construction and commissioning of which is detailed in appendix B.

The MCA we employed is capable of registering pulse heights with a 16 bit resolution and times of arrival with a resolution of 8 ns. We tuned the amplification factor of the incoming signals to make use of most of the range of the MCA's ADCs, resulting in a resolution of 15 ps/bin on the TAC channel and 40 eV/bin on the other. Unluckily the software and drivers shipped with the instrument did not allow us to automate long acquisition processes nor to perform online search of coincidences to perform acquisition monitoring during the course of the experiment, so we had to develop custom drivers. Luckily for what was probably an oversight of the manufacturer the instrument's Linux drivers shipped complete with their source code. This allowed us to reverse-engineer the MCA, task that took over one month and evidenced some critical bugs in the driver implementation that rendered some functionalities

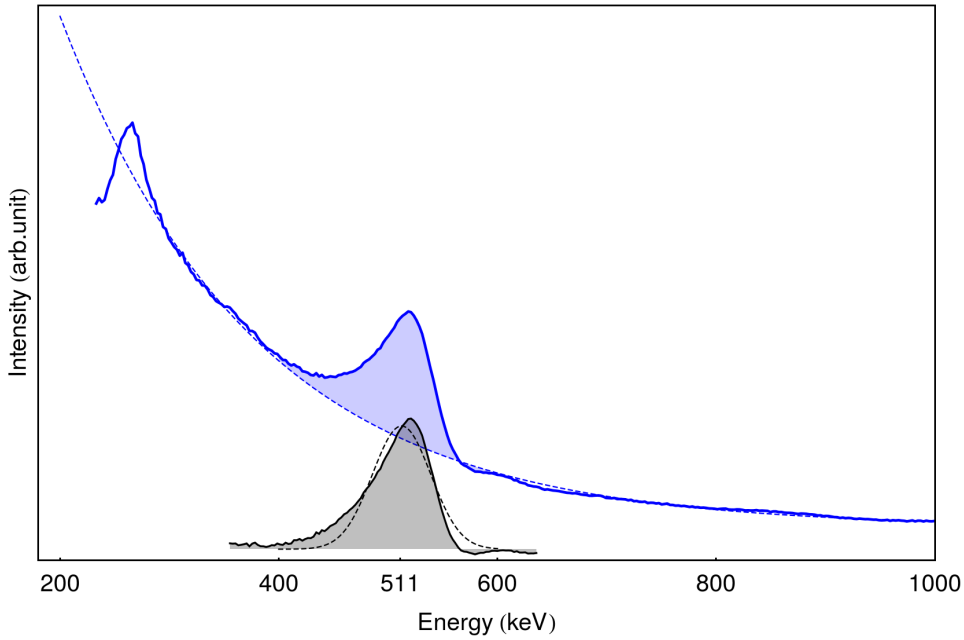


Figure 2.4: In blue the energy spectrum acquired by our NaI(Tl) detectors, in dashed blue an exponential fit to the tails of the 511 keV peak, in black the fit-subtracted spectrum and in dashed black the best gaussian fit to the fit subtracted spectrum which has been employed to calibrate the energy scale. The background subtraction is essential to get an unbiased calibration since, due to the heavy shielding placed in front of the NaI(Tl) detectors, the 2γ peak is heavily suppressed. Around 300 keV the Compton edge is clearly visible

of the MCA non functioning and one critical bug in the MCA firmware that due to an unchecked overflow made the MCA start to produce random data whenever operated continuously for more than 2^{47}ns (a little over 1 day and 15 hours). This means that if we trusted blindly the software shipped with the MCA we would have spent potentially weeks or months acquiring random noise to realize the problem only at the end of the acquisition, which highlights the importance of online data analysis when performing complex experiments that involve many potential failing links in the acquisition chain. It must be sadly noted that, despite our detailed reports, the manufacturer refused to acknowledge the defect so we had to set up our software to reset the MCA every time it approached the clock counter epoch.

There are two control operations needed to run the experiment. The first is the regulation of the MCP bias voltage; this control operation has one strict requirement: at no time the MCP bias voltage should ramp faster than

900 mV/s. We enforced this limitation both on the firmware and the hardware side by means of the custom built control electronics, detailed in appendix B. The second control operation consists in fixing the sample's temperature; this is done via a commercial readout box for diode-based thermal probes that provides closed loop control of a resistive heater to set the sample temperature. The box provides an RS232 serial interface that we cabled directly to the control computer. Both the voltage and the temperature setting interfaces provide also readout capabilities that allow the control computer to verify above the set value, the actual value of the two parameters.

2.4 Measuring for sixteen months

Due to the extremely low rate of the ToF signal, which in our apparatus typically ranges from 6 mBq to 18 mBq, long acquisitions are required to collect enough statistic to form an energy spectrum accurate enough to provide informations on the positronium physics.

We employed a single computer, Melchior, to perform the experiment control, data acquisition and online analysis. Melchior is a screenless box managed though the SSH protocol and accessed for the experiment operation through an HTTPs-based web interface. Another similar computer, Balthazar, having a different main destination, has been set up to perform periodic backup of the data acquired from the PRINCEPS experiment to reduce the risk of catastrophic data loss. Similarly data is copied away from Melchior to perform offline data analysis operations. The Melchior server runs on a vanilla Debian 9 operating system and its web interface is based on Apache 2. All of the software employed to run the experiment is custom built to suit the task.

The organization of the software running Melchior can be, for lack of a better term, described as a swarm intelligence with specialized agents; that is instead of implementing a single program that manages the entire control/acquisition operations required we implement a series of smaller resilient programs, hereafter referenced as daemons, each of which is performing a specific task. We set up daemons so that they can communicate with each other and design their behavior so that, even though no daemon exists that delineates in its code the general operation strategy, the global strategy will emerge from the interaction of daemons.

Giving a complete description of the PRINCEPS daemon suite would exceed the scope of this chapter, so I'll restrict myself to giving a single example of how a global strategy can emerge from the daemon interaction. A daemon has been given the task of regulating the MCP voltage and another the task of regulating the Cryostat temperature. In non optimal vacuum conditions, if the sample is kept for long times at cryogenic temperatures, a

layer of ice might form on its surface[2]. Therefore, whenever we cause the temperature of the sample to cross the threshold of 150 K we incur the risk of evaporating ice in proximity of the MCP detectors; as a precaution against the risk of discharges that could damage the detectors, the voltage of the MCPs should be lowered whenever the sample is reheated. In a monolithic program this would have been likely implemented as a change to the MCP voltage, followed by a suitable pause time, followed by the command to set the cryostat temperature. In our architecture the cryostat control daemon publishes ahead of time its intended alterations to the temperature of the sample and follows suite to apply them to the sample. The MCP control daemon is given the responsibility of checking the cryostat schedule and when necessary lower the voltage to protect the MCPs. In this structure the cryostat daemon can be given commands without considering the consequence they could have on the MCPs since all necessary measures will emerge naturally from the system interactions; this structure simplifies the addition of interlocks since they can be implemented at the level of a single daemon and are guaranteed to behave well on the global level.

Performing long measurements without an operator's oversight requires extremely reliable software. The design rules for daemons that we gave ourselves require that it must be possible to terminate or kill any of them at any time, then restart it and the new instance must be able to resume the required operations almost seamlessly. Second requirement is that every daemon must have an externally checkable heartbeat function that should be rendered inoperative by daemon entering any deadlock condition or being stuck waiting for an external input. Daemons are periodically launched by the operating system, even when a previous instance of the daemon is still running. When at startup a daemon detects a previous instance of itself it checks the heartbeat of the previous instance and quit as soon as the heartbeat is detected; if no heartbeat is detected within 15 s the new instance is required to kill the previous one and take its place. Given the structure of the software it is also possible to make it resilient to abrupt power cuts making it capable of resuming the measurement as soon as the power is connected again. We opted to avoid this since we do not have automated a diagnostic on the vacuum present in the chamber (it is scheduled within the next upgrade) so Melchior cannot grant that the vacuum system is still in working order after a power cut.

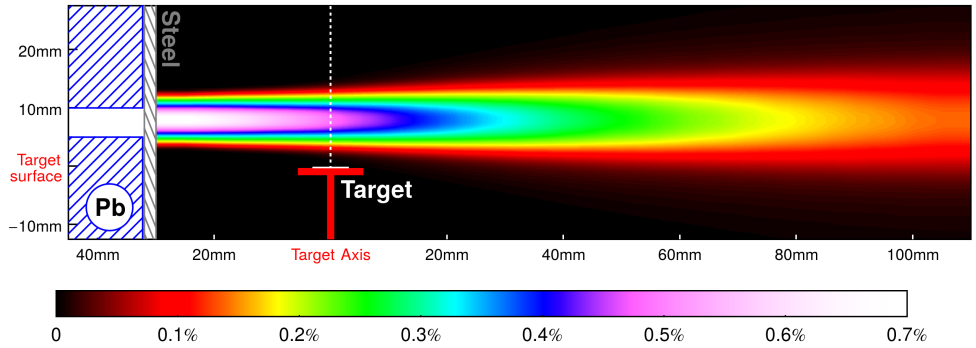
On top of everything a daemon provides a remote interface running over the Telegram protocol that provides control functionalities, periodic recapitulation of the apparatus status and instant notification in case an interlock is triggered. With this setup we have been able to measure almost without interruption for

[2]Although we will always operate the apparatus so to avoid this occurrence at all costs, the control software should be resistant to non optimal operation.

16 months.

2.5 Samples tested

For all of the measurement in this campaign we set the slit's opening to 5 mm and its center 7.5 mm, which results in the lower lid of the slit to be placed 5 mm and its upper lip 10 mm above the target surface. This, naturally, does not mean that every decay that we are going to detect will be situated between these two extremes. The acceptance (defined as the density of probability per unit of volume that a particle decaying in a set position in the chamber will produce a γ ray that will traverse the open region of the slit and hit a detector) is shown in figure 2.5



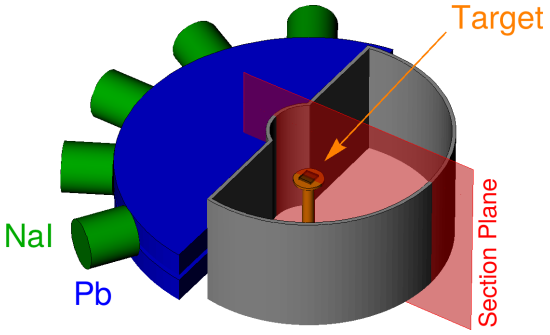


Figure 2.5: Above: geometric acceptance (as defined in the text) for the employed slit. Being the acceptance a three-dimensional function its value is shown in section, employing the plane shown in the diagram on the side.

Before anything else we performed a prompt measurement by placing a non etched, not oxidized monocrystalline silicon chip on the sample holder and implanting positrons into it. We measured the resulting ToF spectrum when implanting positrons at an energy of 7 keV collecting 26.3 k ToF, 26.6 M NaI(Tl) events. We used this first prompt to perform online data analysis and monitor the collected data as the measurement was still ongoing. At the end of the entire measurement campaign, on the first of February 2018, we initiated the acquisition of a second prompt at an energy of 11 keV the measurement of which is still ongoing. The 11 keV prompt is necessary to

Shorthand identifier	Sample	Temperature	Implantation Energy	ToF Samples	NaI(Tl) Samples
Prompt7	Si	300 K	7 keV	26.3 k	26.6 M
Prompt11	Si	300 K	11 keV	*Ongoing*	*Ongoing*
300K	A	300 K	7 keV	50.1 k	48.4 M
20K	A	20 K	7 keV	61.2 k	65.9 M
11keV	A	20 K	11 keV	68.2 k	100.1 M
Thin	B	20 K	7 keV	67.5 k	153.6 M
250K	B	250 K	7 keV	64.0 k	158.5 M

Table 2.1: Color-coded shorthand identifiers for each measurement and measured statistics. The color scheme presented here will be employed across all of the plotted data coming from the PRINCEPS experiment.

properly analyze the measurements performed at a higher energy; since this data is not yet available we are forced to use, in this work, the 7 keV prompt to analyze all data and defer the full data analysis to a future work.

We then measured the ToF spectrum of two samples that we will name sample *A* and *B*. The two samples were generated with the electrochemical etching process described in [54] to produce the sample there named respectively #2 and #0. The two samples are differentiated mainly by their mean channel diameter that is expected to be about 13 nm in sample *A* and 6 nm in sample *B*,*Mariazzi1*.

We measured sample A at a temperature of 300 K with an implantation energy of 7 keV and a temperature of 20 K with implantation energies of 7 keV and 11 keV. Then we measured sample B with an implantation energy of 7 keV at a temperature of 20 K and 250 K. A recapitulation of the measurements is shown in table 2.1

Measurement performed at room temperature have been conducted without interruption (with the exception of the periodic reset of the MCA due to the firmware overflow bug). The measurement performed at 250 K was paused every 28 hours to heat the sample to a temperature of 300 K for about 15 minutes to prevent the condensation of residual contaminants on the sample. The measurements performed at 20 K was paused every 8 hours to heat the sample to a temperature of 300 K for about 7 minutes to prevent the formation

of ice on the NCP surface. We are operating the experiment with a residual gas pressure below $1 \cdot 10^{-9}$mbar which we expect to take several days to build up a significant layer of ice on the sample; we confirmed the absence of ice formation by finding no difference in rates or spectrum of data acquired shortly before and shortly after reheat cycles. Figure 2.6 shows the temperature of the sample as a function of time during a heating cycle in a 20 K measurement.

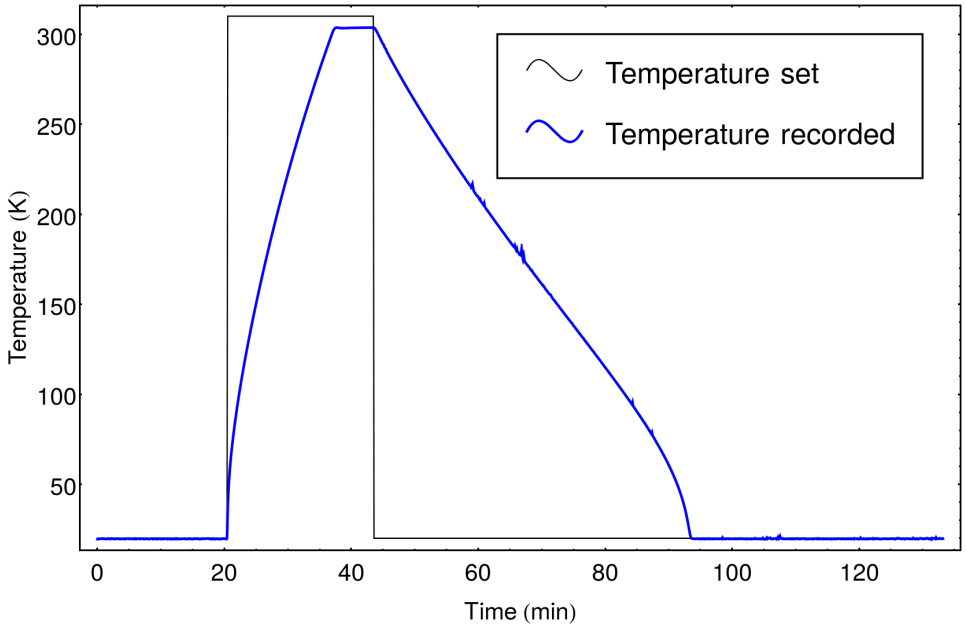


Figure 2.6: A re-heating cycle of the sample during a 20 K measurement. In blue the sample temperature measured obtained my measuring the drift of the V-I response of a calibrated diode. The discrepancy between the measured high temperature and the requested temperature is due to the heating element providing the thermostasis being connected in a closed loop onto an analog feedback whose precision decreases with an increasing temperature.

2.6 Data analisys

Raw data from the PRINCEPS experiment comes as two long series of timestamped pulse heights, namely the ToF signals and the NaI(Tl) signals. We calibrated the ToF ADC scale by employing a function generator and a delay line to inject carefully prepared signals with a known delay into the ToF acquisition chain. We calibrated the NaI(Tl) ADC scale by placing an ^{22}Na radioactive source within the slit and measuring the 511 keV peak position

after fitting and subtracting its background as shown in figure 2.4.

The first step of the data analysis procedure is to form coincidences, that is: we enlist all pairs of samples (one sample taken from the ToF list and one sample taken from the NaI(Tl) list) for which the timestamp difference between the two samples is less than 1 µs. The choice of a window is extremely loose since events in coincidence usually are recorded within a 100 ns time window and rates are low enough that even windows of a few tens of microseconds produce very few spurious coincidences. It is possible given this definition that a single event will be enlisted in more than one coincidence, this happens typically in less than 50 cases per million coincidences which is largely above the level of accuracy required by our experiment, therefore we decided to not operate any further selection and, in these cases, accepted the multiple coincidences as all valid.

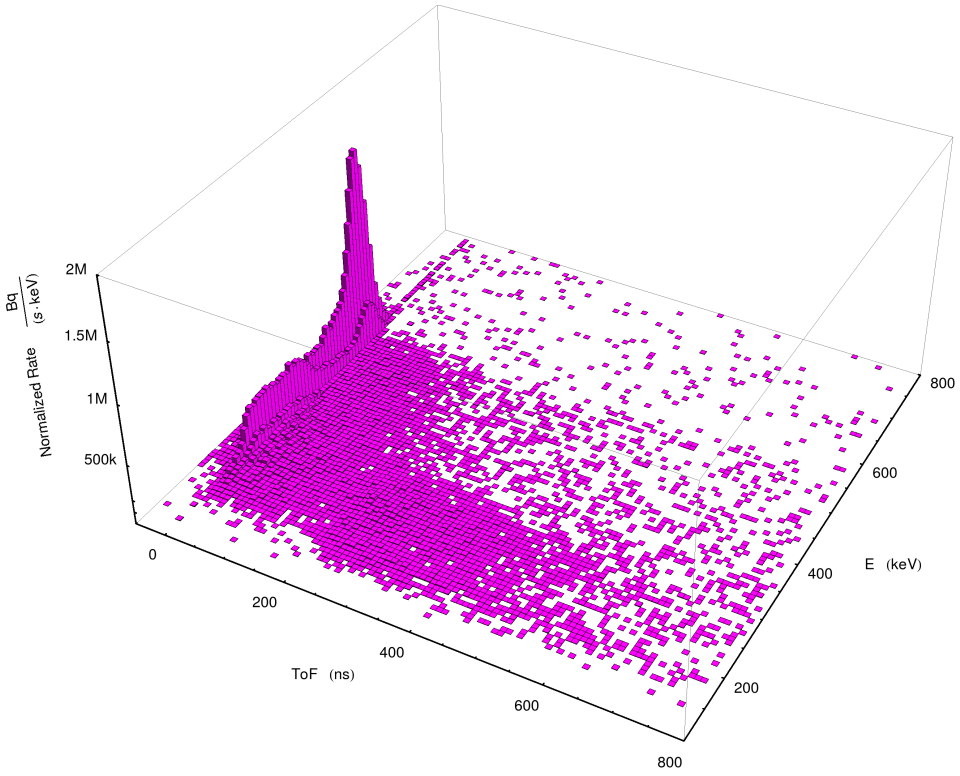


Figure 2.7: Event distribution of the **20K** measurement coincidences as a function of the recorded γ energy and the positronium time of flight.

We can now get a general idea of the acquired data by plotting all of the gathered coincidences in a 3D histogram as we did in figure 2.7 with the data gathered from the **20K** measurement. In the figure it is possible to distinguish

the prompt peak at low timings ($< 20\,\text{ns}$) and high energies ($> 400\,\text{keV}$). We expect the thermalized positronium to generate a signal in the, now almost empty, region at slightly higher timings ($< 250\,\text{ns}$) and high energies ($> 350\,\text{keV}$). At high timings ($> 100\,\text{ns}$) and low energies ($< 350\,\text{keV}$) we see a detached "island" of signal coming from backscattered positrons annihilating on the chamber walls. We need to exclude this contribution, which is orders of magnitude larger than the positronium contribution, to obtain a clean signal, so the next passage in the data analysis will consist in excluding all of the coincidences in which the NaI(Tl) recorded a γ energy greater the $350\,\text{keV}$. The effectiveness of this operation can be see also when we plot the ToF signal recorded in the coincidences against the difference between the timestamps of the two signals. In an ideal world we would expect the resulting points to lie on a one dimensional manifold; this is not the case for us since the jitter on the ToF line is unmanaged, still the backscattered positron contribution can be clearly seen in the plot as a separate distribution of points that is completely erased when the low energy cut is applied (see figure 2.8).

We then flatten the distribution maintaining only the ToF data and throwing away the timestamps and the energies. What we obtain, due to the nature of the ADC of being capable of delivering only values taken from a finite and reasonably small ($65.5\,\text{k}$) set of possible value is already in the form of an histogram. Although the binning is too fine to be used to display data we will manipulate it without rebinning as long as we can since rebinning operations in general destroy information.

The next operation required is to perform a background subtraction, that is we want to eliminate the contribution to the signal coming from the source itself, from losses in the transport, from positrons annihilating into the target and from backscattered positrons annihilating onto the chamber walls. First of all we subtracted any plateau noise present in the data. To do so for each single histogram we compute the average bin content between ToF values of $600\,\text{ns}$ and $800\,\text{ns}$ and subtract the result from the content of each bin in the histogram. Then we use the prompt peak to normalize all of the histograms. To do so we sum the content of the bins in the range $\pm 5\,\text{ns}$ for the ToF value, then divide the content of each bin in the histogram by the result.

Before continuing the exposition let us present the histograms resulting from the plateau removal (see figures 2.9, 2.10 and 2.11). To allow for readability of the graphs by the human eye we rebinned the data into bins with a size comparable with the peak width. The peak width at half maximum is $2.5\,\text{ns}$ in all of the measurements and $5\,\text{ns}$ at 10% height, we therefore chose to employ $4\,\text{ns}$ wide bins for the visualization. The region of the plots that is significant for the subsequent data analysis is the region comprised between $45\,\text{ns}$ and $200\,\text{ns}$ of the horizontal scale. Due to the high statistical noise, the presence and shape of the signal (particularly in figure 2.11) is not always easily readable by eye; we therefore included an inset that shows the spectra

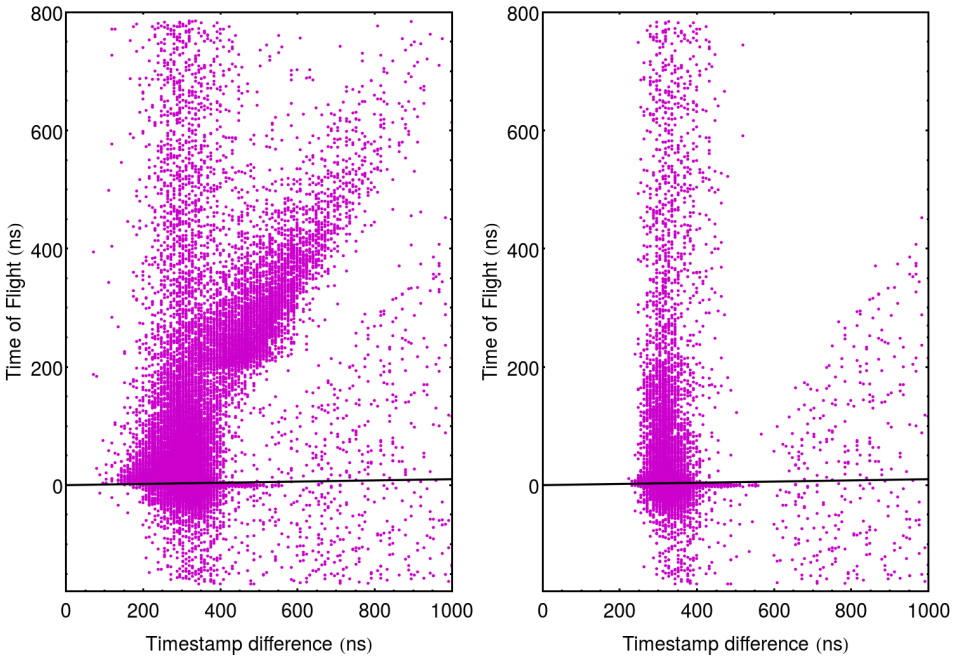


Figure 2.8: Scatterplot of the difference in time of arrival for events coming from the two channels of the MCA (one being the ToF signal and the other the γ ray energy) against the actual ToF signal. On the left all of the recorded events are shown, on the right only events whose γ ray energy exceeds 350 keV. We can see in the left scatterplot two main features: one with constant timestamp difference, the other displaying a linear dependency. Since the NaI(Tl) crystals produce the stop signal we expect the timestamp difference to be roughly constant and the linear-dependent component to be caused by electronic crosstalk. We can clearly see the linearly-dependent feature being completely cleared out by the cut operation.

in the most relevant region after they have been smeared by averaging through a 64 ns-wide rectangular moving window. Although the **11keV** measurement should be compared to the prompt peak **Prompt11**, due to the current unavailability of such prompt we had to plot it against **Prompt7**. From the comparison of the data presented in the different plots we can see that the signal is much more prevalent in the **300K** and the **20K** measurements than in the **Thin** and **250K** runs. This is expected in the framework of chapter 1 due to the 3γ fraction decreasing as the channels get thinner, since more interactions are required to exit the NCP resulting in a greater fraction of positronium annihilating onto the walls of the channels during thermalisation.

In figure 2.12 we show the spectra for the measurements **300K**, **20K**, **Thin** and **250K** as presented in the insets of figures 2.9 and 2.11 (that is, averaged over the 64 ns moving window) to which the prompt **Prompt7** had been subtracted. We excluded from this representation the **11keV** spectrum since we do not have the necessary **Prompt11** to perform the appropriate subtraction. Bearing in mid that lower energy positronium will appear in the right side of the spectrum, we can observe that, as expected, the **20K** measurement displays more cold positronium than the **300K** measurement. The same can be predicated with the measurements performed in thinner-channeled samples with the **Thin** measurement featuring more cold positronium than the **250K** measurement. We can further observe that the **250K** measurement decreases heavily after the 150 ns mark, similarly to the **300K** measurement. Further characteristics are difficult to assess through naked eye observation, so we will proceed with the rest of the analysis.

We proceed in subtracting the background that we measured by taking the measurements named *Prompt7* and *Prompt11*. Now a strategic decision must be taken: we could either subtract directly the measured prompt or we could fit it with a suitable empiric model and subtract the model from the measurement histogram. The drawback of the first method is that it will greatly increase the result uncertainty. On the contrary the latter method will bias the data in case some small feature is present in the prompt data that cannot be described by the employed empiric model. Basing on the shape of the prompt histograms and an expected absence of small features, as dictated by experience, we opted for the latter method. We then fitted the two prompt curves in the $30\,\text{ns} \div 240\,\text{ns}$ range with a cubic polynomial. The range was chosen to include all and only the data which is relevant to the construction of the positronium energy spectrum. The fitted function is then subtracted from all of the measurements performed at the same energy at which the prompt was acquired.

After the background subtraction we are ready to construct the Ps energy distribution. We will proceed similarly to[103]. First of all we relabel the bins by transforming ToF value into positronium energies. The slit not having a delta shaped acceptance we will have to take an average value for the

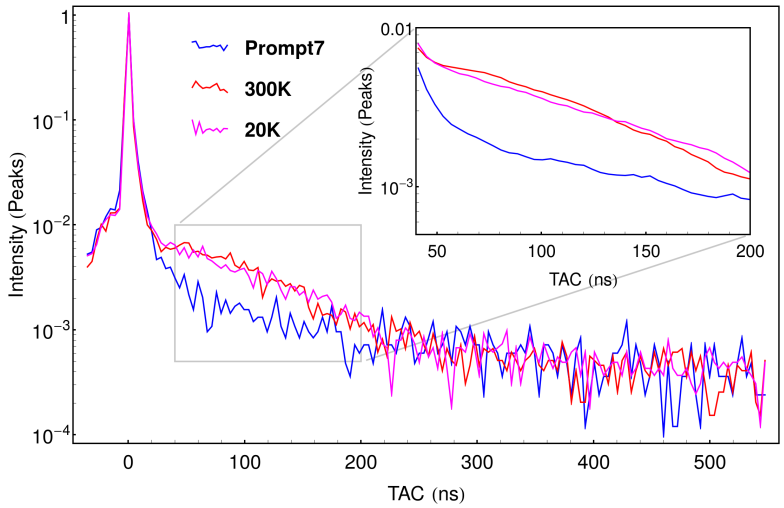


Figure 2.9: 7 keV ToF prompt spectrum measured on a silicon chip (blue) superimposed to the **300K** (red) and **20K** (magenta) spectra. In the inset the same spectra are presented averaged over a moving window of width 64 ns.

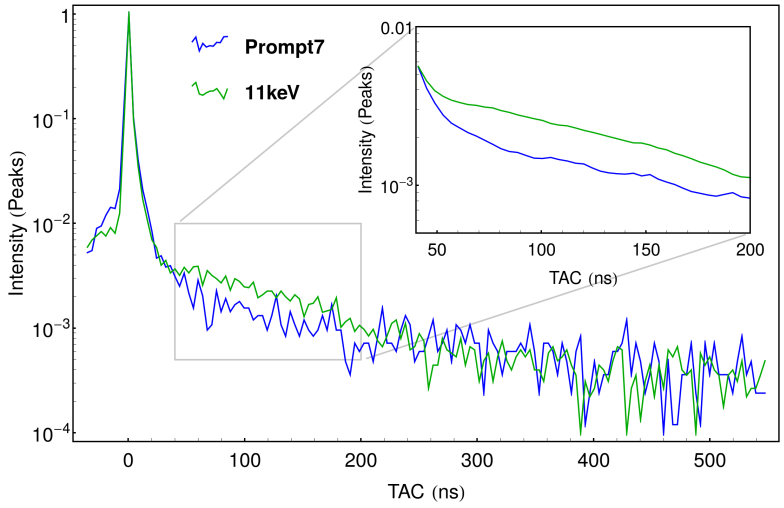


Figure 2.10: 7 keV ToF prompt spectrum measured on a silicon chip (blue) superimposed to the **11keV** (green) spectrum, similarly to image 2.9. The actual prompt that should be subtracted from the spectrum is indeed the 11 keV one, whose measurement is still ongoing.

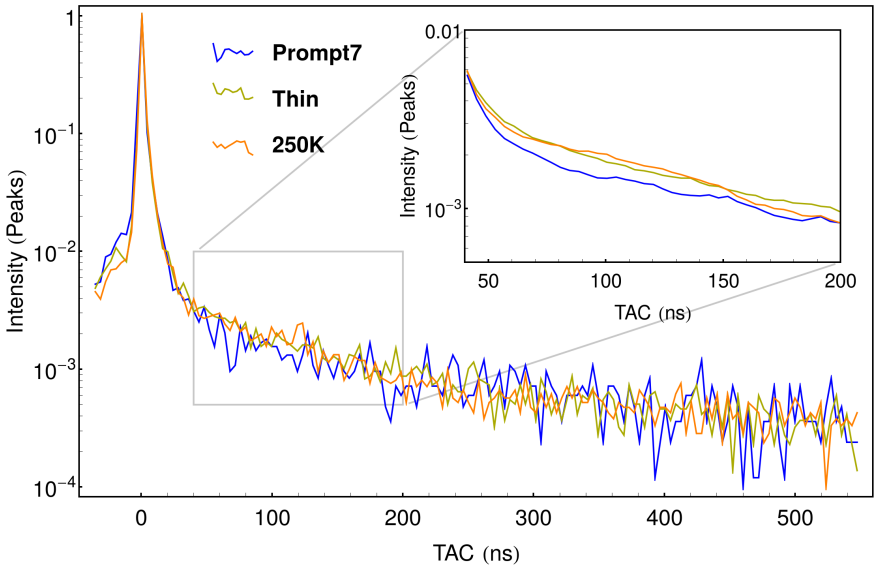


Figure 2.11: 7 keV ToF prompt spectrum measured on a silicon chip (blue) superimposed to the **Thin** (ocra) and **250** (orange) spectra, similarly to image 2.9.

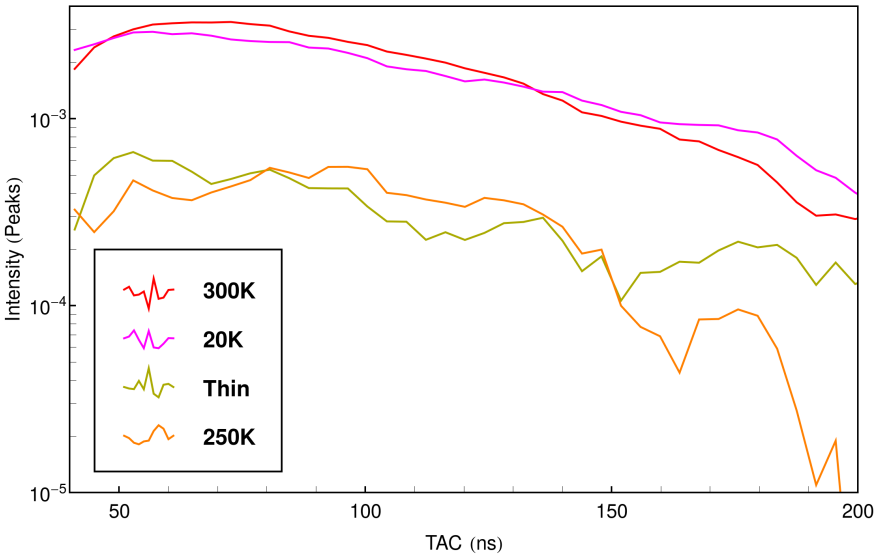


Figure 2.12: **300K**, **20K**, **Thin** and **250K** spectra, averaged over a 64 ns moving window, to which the prompt **Prompt7** has been subtracted.

positronium kinetic energy, an excellent estimate of which is given by the naïve model obtained by writing the positronium speed as:

$$v_\perp \;=\; \frac{d}{\mathrm{ToF}} \;=\; \frac{7.5\,\mathrm{mm}}{\mathrm{ToF}} \tag{2.1}$$

in which the distance d has been taken as the distance between the plane that cuts the slit in half and the detector surface; to then write the positronium kinetic energy as:

$$K_\perp \;=\; \frac{1}{2}mv_\perp^2 \;=\; \frac{4.1048\cdot 10^{-16}\mathrm{eVs}^2}{(\mathrm{ToF})^2} \tag{2.2}$$

with m the positronium mass. Due to the almost perfectly symmetrical geometry of the slit this estimate is actually accurate up to $8\cdot 10^{-6}$.

Then since the relabeling of the histogram axes was not performed using an affine transformation we need to compensate the change in the histogram bin widths by multiplying by the Jacobian of the performed transformation. We will therefore multiply the content of each bin by the weight function:

$$w^{**}(\mathrm{ToF}) \;=\; (\mathrm{ToF})^3 \tag{2.3}$$

This function is not complete, we need to compensate for the effect given by the fact that particles traveling at higher speeds spend less time within the acceptance volume of the slit therefore making them less likely to be detected. Therefore:

$$w^{*}(\mathrm{ToF}) \;=\; (\mathrm{ToF})^2 \tag{2.4}$$

We still need to compensate for the finite life of the positronium in vacuum that biases the measurement by reducing the detected signal as the value of the ToF increases. The final weight function will therefore be:

$$w(\mathrm{ToF}) \;=\; (\mathrm{ToF})^2\cdot e^{\mathrm{ToF}/\tau} \tag{2.5}$$

where $\tau \approx 142.05\,\mathrm{ns}$ is the positronium lifetime in vacuum. Figure 2.13 shows the shape of the weight function plotted as a function of the positronium energy, thus highlighting the main limitation of our apparatus: the slope of the weight function increases greatly as the positronium kinetic energy approaches zero, therefore any uncertainty in the measurement of the ToF for Ps energies below $\approx 18\,\mathrm{meV}$ will be unsustainably amplified in the final result thus limiting the minimum energy to which we can extend our spectra.

We mentioned before that we avoided rebinning the histogram resulting naturally from the granularity of the ADC data during the entire manipulation; this is advisable since it introduces no bias due to rebinning but, at the same time, would yield plots that are difficult to read. In the proposed plots we applied a final rebinning that we calibrated to optimize the clarity of the resulting plot. We employed bins of variable width, ranging from 1.2 meV to 4.5 meV with the larger bins located at the center of the spectrum (around 0.15 eV) and the thinner ones at the sides. We present all of the measured spectra (let's name them $\mathcal{M}(E)$) in figures 2.17 to 2.21 located at the end of the following paragraph so that they can be accompanied with fits obtained from the model described in chapter 1; we nonetheless invite the reader to peek forward and familiarize with the data as we briefly describe it.

The first characteristic that we notice on all of the spectra is that all of them display some thermalisation of the positronium, the energy distribution being always comprised mostly below 0.1 eV. Secondly if we compare data gathered from sample A and sample B we notice that the measurements in sample B display in general lower energies than their counterparts performed on sample A (although we are forced to compare the **300K** measurement with the **250K**). Moreover measurements on sample B show lower statistics than the measurements performed on sample A, this is expected both the literature [54] and from the model presented in chapter 1. According to our model thinner channels require more interactions to be escaped, which in turn increases the probability of wall annihilation before escape thus reducing the amount of produced positronium. If we compare the measurements performed on sample A we clearly see the increase in the positronium cooling caused by the employment of a colder sample. The **300K** measurement shows a plot whose maximum could be assumed to be close to the minimum energy detectable by PRINCEPS, while for the **20K** and **11keV** measurements the peak likely lies below the visibility threshold. This is, indeed, expected for colder samples since the expected mean energy for a 300 K lies above the limit of 15 mK whereas for a thermal distribution with an energy of 20 K it lies so close to zero as to be indistinguishable from it in our current binning. We will therefore have rely on fitting against the model if we want to determine the position of the peak of spectra measured at 20 K. If we compare the measurements performed at 20 K we can see that, as expected both the reduction of the channel size and the employment of a higher implantation energy yield an overall colder positronium spectrum; this once more is in accordance both with the trends observed experimentally[54, 55] and with the predictions given by our model. If we compare the **11keV** and the **Thin** measurements we observe that the **Thin** measurements presents a slightly colder spectrum than the **11keV** measurement; this inequality relies on several components of the thermalisation process and cannot be easily predicted without running a simulation, nonetheless the model of chapter 1 predicts it correctly.

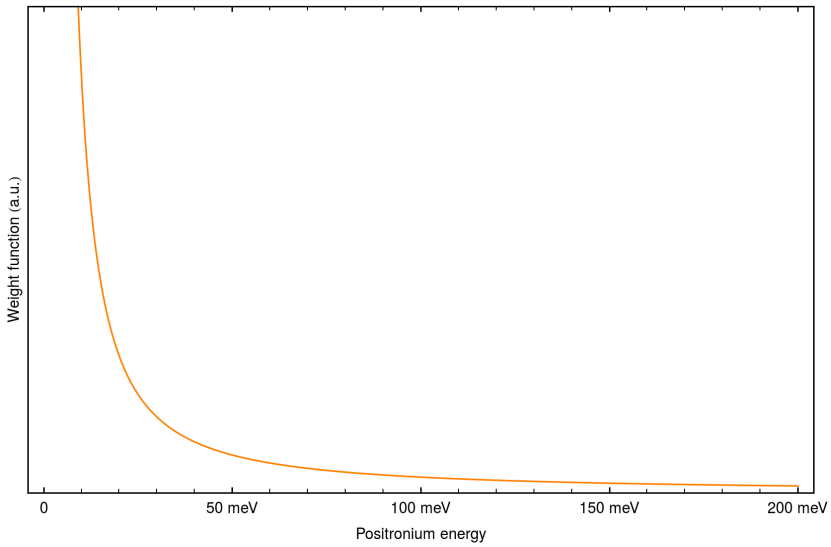


Figure 2.13: Weight function by which positronium kinetic energy spectra must be scaled. It can be noted that for energies below $\approx$ 18 meV the slope of the weight function becomes extremely steep.

The final uncertainty on the spectrum is given mainly by two factors: the finite amount of events collected and the non null width of the slit acceptance distribution. The former can be easily estimated from Poisson's distribution variance. Estimating the effects of the latter is no easy task. The effects of the slit's finite width can be described, in first approximation, as the actual energy spectrum being convoluted with an appropriate distribution. In this scheme the effect could be, in principle, removed by deconvolving the measured spectrum; nonetheless deconvolution techniques are intrinsically numerically unstable therefore, considering the low signal to noise ratio in our measurements, shall this techniques fail we would have no way to recognize it.

It is instead easier to mimic the effects of the finite slit onto the simulated data and to compare the result with the experimental data.

2.7 Applying the model

As anticipated, to compare the experimental data to the simulations we simulate the effects of the finite slit width onto the simulated Ps spectra and compare the resulting distributions with the measurements.

To do so we run a Monte Carlo simulation in which Ps atoms are emitted from a point on the surface of the nanochanneled plate drawn with a distribution that mimics the shape and size of the PRINCEPS positron beam

obtaining from it an energy spectrum $\mathcal{K}(E)$. We then run a Monte Carlo simulation to simulate the ToF data that the experiment would collect in this situation. Emission angles for the positrons are deduced from the geometric component of the thermalisation simulation. Emitted Ps atoms will be made to decay in flight after an exponentially-distributed random time, from the annihilation site three gamma rays are emitted and, shall their trajectory match the slit opening and hit a detector, an event is generated. In practice it is mathematically equivalent and much more efficient to count all of the annihilation events weighting them with the acceptance function a section of which is displayed in figure 2.5. The ToF data generated in this process are then gathered in an histogram that is fed into the same algorithm employed to analyze real ToF data. Although the simulation does not produce a prompt peak, the code analysis correctly detects its absence and performs no peak subtraction.

In figure 2.14 we show, both in its original form and how we expect PRINCEPS to see it, a spectrum simulated with an effective $M = 8\,\text{amu}$, fixed wall annihilation of $1/p = 200000$, 11 nm-wide dendritic channels a sample temperature of 300 K an implantation energy 7 keV and a wall interaction scattering with $q = 3$. These parameters have been chosen to be distant from the ones used for the plots displayed ad the end of the chapter with the purpose to show that different sets of parameters tend to yield similarly-shaped spectra.

We will try to extract an experimental value for the effective mass M parameter which, as defined in the model described in chapter 1, quantifies the rate at which positronium loses its energy as it interacts with the channel walls. To do so we will simulate spectra employing the simulator described in chapter 1, modify them to account for the response of the PRINCEPS apparatus. Then we normalize the simulated spectrum and the measured one we want to compare it to so that:

$$\int_{15\,\text{meV}}^{\infty} \mathcal{K}^*(E) \;=\; \int_{15\,\text{meV}}^{\infty} \mathcal{M}(E) \;=\; 1 \tag{2.6}$$

where $\mathcal{K}^*(E)$ is the simulated spectrum, $\mathcal{M}(E)$ is the measured one and 15 meV is PRINCEPS' limit energy at which data is still significant. To evaluate how close the simulated spectrum is to the measured spectrum we compute the integral:

$$Q = \int_{15\,\text{meV}}^{\infty} (\,\mathcal{K}^*(E) - \mathcal{M}(E)\,)^2\, dE \tag{2.7}$$

In which the integral is performed over linear interpolation of the discretely sampled functions $\mathcal{K}^*(E)$ and $\mathcal{M}(E)$. The value of Q is a quality factor akin to

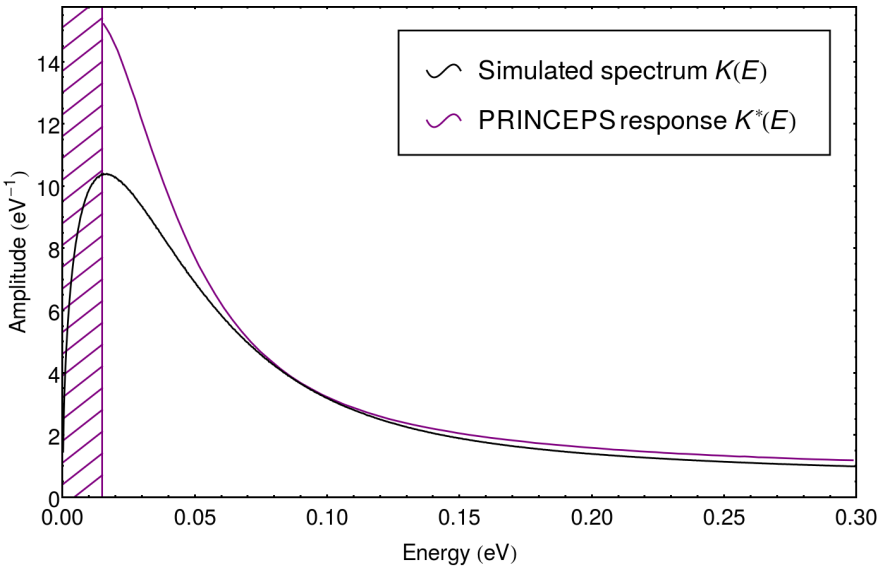


Figure 2.14: A positronium energy spectrum simulated through the procedures detailed in chapter 1 (in black) and the same spectrum as we expect PRINCEPS to record it (in purple). The region excluded from the purple plot correspond to the blind spot of PRINCEPS caused by the weight function becoming too steep (see the end of section 2.6). Two main characteristics differentiate the two plotted spectra. The first is that, due to the blind spot in the PRINCEPS response, normalizing $\mathcal{K}^*(E)$ to have a unitary area (i.e.: rendering it a PDF) yields a higher curve than $\mathcal{K}(E)$. Secondly, even if the two spectra are normalized to lay one on top of the other the PRINCEPS response in, generally, more low-energy-heavy due to faster positronium exiting the target with a non null angle with respect to the vertical being interpreted as slower positronium emitted orthogonally from the target surface.

χ^2, in which lower values of Q mark simulated spectra closer to the measured ones and larger Qs mark worst-fitting spectra. We minimize Q by repeating the simulation for different values of the effective mass, the effective mass yielding the lowest value of Q is the result of the fit procedure.

To provide an estimator of the uncertainty over the final value of M, we used the experimental data to produce similarly-distributed spectra and quantified the stability of the fit procedure over these spectra. We did so by employing the particle count h_i present in the ith bin as the mean of a Poisson distribution that we then employ to draw the particle count of the ith bin in the generated histogram. We then process the newly generated histogram in the same way we processed the experimental data obtaining, thus, a new value for the effective mass M. By repeating the procedure 128 times we obtain a population of values for M whose variance is given by statistical fluctuations similar in magnitude to those we expect from the finite statistics gathered from the experiment. We will employ the square root of the variance of the generated M distribution as an estimator of the uncertainty over the M reconstructed from the experimental data.

2.8 Results and Interpretation

If the completely classical model were capable of describing satisfactorily the thermalisation process in NCPs we would expect the effective mass resulting from the the fit procedure to be compatible across all five measurements. Moreover, due to the chemical composition of the NCP we do not expect M to be less than that of an oxygen atom (16 amu); we can use these two criteria to determine if the model formulated in chapter 1, which was calibrated on experimental data gathered mostly with samples held at room temperature, holds on a wider temperature range.

First of all we attempted the reconstruction by employing a constant wall annihilation factor $1/p = 100000$, the wall scattering distribution obtained by employing $q = 3$ (as defined in section 1.9) and the channel radii, sample temperatures, implantation energies relative to the specific condition each measurement had been performed in. Figure 2.15 shows the result of the M fit for all the five measurements performed. Although we know the radius of the nanoscopic channels in our samples we present here the result of the same procedure repeated for different values of the channel radius so that the resulting landscape can be compared to that of figures 1.18 and 1.23. Of the five measured spectra, only the **300K** measurement is capable of retrieving a value compatible with our expectations, the **250K** measurement comes close to do so while all of the others indicate that the model, as employed here, is incapable of justifying the shape of the recorded spectra. Notably the effective mass seems to depend only on the sample's temperature, with the cold spectra

(that of measurements **20K**, **11keV** and **Thin**) showing compatible Ms for different implantation energies and channel geometry.

Before drawing our conclusions there is one question that we need to address. How does choice of the p parameter influence these results? Moreover, is the model presented in section 1.11, which introduced a dependency of p on the positronium temperature, capable of better explaining the experimental data? We repeated the entire procedure employing the temperature-dependent p model. Figure 2.16 shows the resulting landscape in the same format as figure 2.15. The most striking difference introduced by the temperature-dependent p model is that the measurements performed onto sample B (**Thin** and **250K**), albeit being vexed by a high experimental uncertainty, are now compatible with physically plausible values of M, while the low temperature measurements of sample A still yield values for M only slightly above 1 amu.

It could be argued, since the effect of modifying the p parameter was to increase the reconstructed effective mass M on all of the measured spectra that the reason why the model yields unrealistically low values for the effective mass of two out of the five measurements is that an improper value of p is still being used. The effect of a high p on the spectrum is to depress the portion of positronium which needs more wall interactions to exit the channel, thus increasing the average energy of the positronium that leaves the sample. When the effective mass is fitted through the model a high p parameter will induce a lowering of the reconstructed M, since a portion of the positronium spectrum which underwent a lower number of interactions will have to replace the quenching-depressed high-interaction-count portion. Since a lower value of M will ensure a faster thermalisation, the reconstruction process will opt for it to provide enough cold positronium to fit the experimental data.

Regardless of whether the value of p is taken to be constant or variable, if we run the reconstruction with $p = 0$, unrealistically corresponding to no 2γ annihilations taking place inside of the nanoscopic channels, we expect the highest possible values for the effective mass to be returned by the process. If we do so the **20K** and **11keV** measurements yield respectively $M = 1.4 \pm 0.4$ amu and $M = 1.7 \pm 0.3$ amu , both still incompatible with the expected 16 amu; useless to say not even the coherence of the value of M between measurement is recovered now.

No definitive conclusion should be drawn, of course, until the 11 keV prompt has been measured and the entire data analysis chain has been run again. For the moment, based on the observations reported here, a possible interpretation is that we are effectively seeing the limit of Sauder's model when approaching low temperatures; this is more evident as the temperature of the sample is lowered, since the De Broglie wavelength of positronium can reach higher (the De Broglie wavelength of a 300 K positronium atom is about 5 nm, at 20 K it becomes more than 20 nm). If this interpretation is correct, then the measurements in thinner channels are expected to also

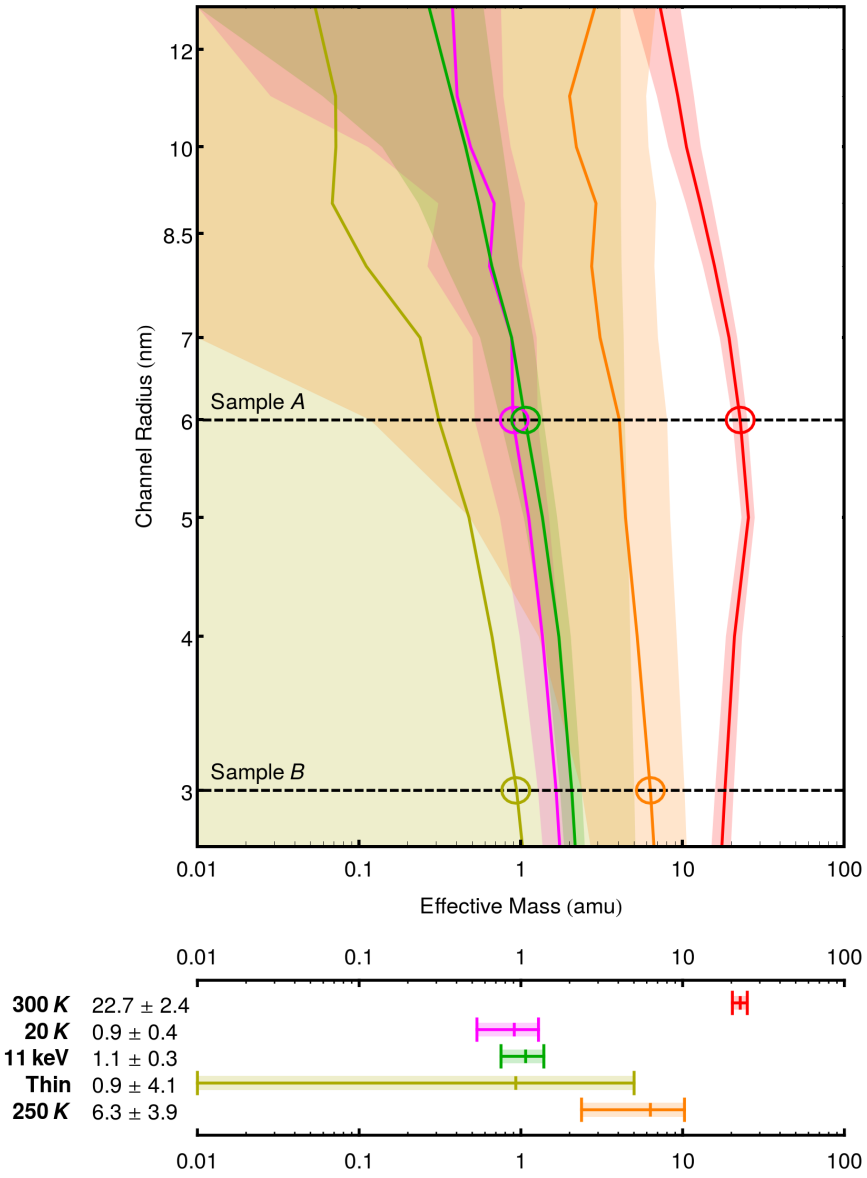


Figure 2.15: Result of the M fit (as described in the text) for different values for the channel radius computed wit $1/p = 100000$. The y axis is labeled as in figure 1.23, the dashed lines represent the mean channel radius we expect the two measured samples to feature. On the bottom the fit result for all of the measurements.

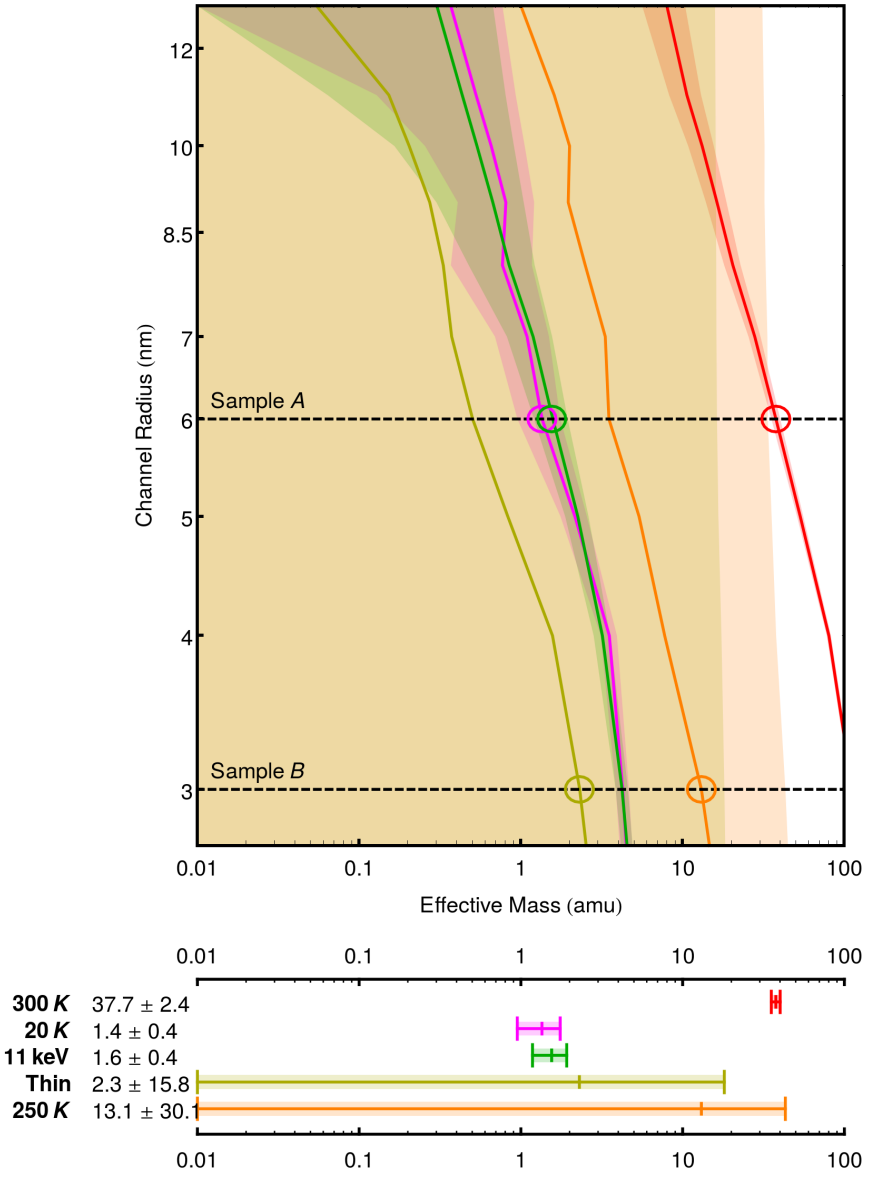


Figure 2.16: Result of the M fit (as described in the text) for different values for the channel radius computed with the temperature-dependent p model proposed in section 1.11, the dashed lines represent the mean channel radius we expect the two measured samples to feature. The y axis is labeled as in figure 1.23. On the bottom the fit result for all of the measurements.

break the model's capability of reconstruction possibly in a more dramatic way; nonetheless this is not necessarily bound to manifest in the form of a low value of the reconstructed effective mass. While the thermalisation physics is expected to transition from a classical to a quantum regime sooner in thinner channels than in larger ones, the increased number of expected wall interactions will grant more leeway to the classical model to increase the effective mass and still be able to predict a large enough amount of thermalized positronium to match the experimental data. Under this interpretation we would expect a larger error bar for the thinner channel measurements (due to a larger span of effective masses granting enough thermalized positronium) larger error bar that we, indeed, observe. Moreover we would expect the classical model to still struggle to explain the behavior thinner-channeled samples which is, again, the case in the experimental data. Finally we would expect the classical model, even when applied to thinner channels, to still perform better on measurements performed on a hotter sample, assertion that, again, is supported by the experimental data.

As said in chapter 1 the $\mathcal{K}(E)$ spectra obtained from the simulator can be fitted with two thermal distributions. As we saw in section 1.10, these two distributions arise from the interplay of the thermalisation-escape processes and do not reflect actual physical quantities; on top of that by fitting a spectrum obtained from a model that we know to describe incorrectly the thermalisation dynamic we derive values that should be taken with a grain of salt to say the least. Nonetheless for the sake of completeness we'll report that the cold fraction of all of the fits, with both models for p yields a temperature that is slightly higher than that of the sample with the discrepancy increasing as the temperature is increased. The three 20 K measurements all yield fitted temperatures for the cold fraction comprised between 21.5 and 23.5 K, the **250K** measurement between 280 and 290 K and the **300K** measurement between 410 and 420 K. The amount of the cold fraction indicated y the fit is highest for the cold measurements (with values ranging from 70 to 90%) and lower for samples with a higher temperature (55-60% for the **250K** measurement and 40-45% for the **300K** measurement). The temperature of the hot fraction varies, depending on the specific fit between 900 and 14 000 K with the high values being mostly due, as we said in chapter 1, to the window in which we are able to perform the fit being wide. These numbers can be used effectively to describe the spectrum resulting from the simulation and for comparison with other experimental measurements, still they should be regarded as nothing more than an excellent empiric description.

The next step in this investigation could be to employ one of the proposed quantum thermalisation models[57] to attempt to explain the behavior of cold NCPs. We will end this chapter by presenting all of the measured spectra upon which we superimposed the best fit provided by our model both when assuming $p = 10^{-5}$ and the temperature-dependent p (figures 2.17 to 2.21).

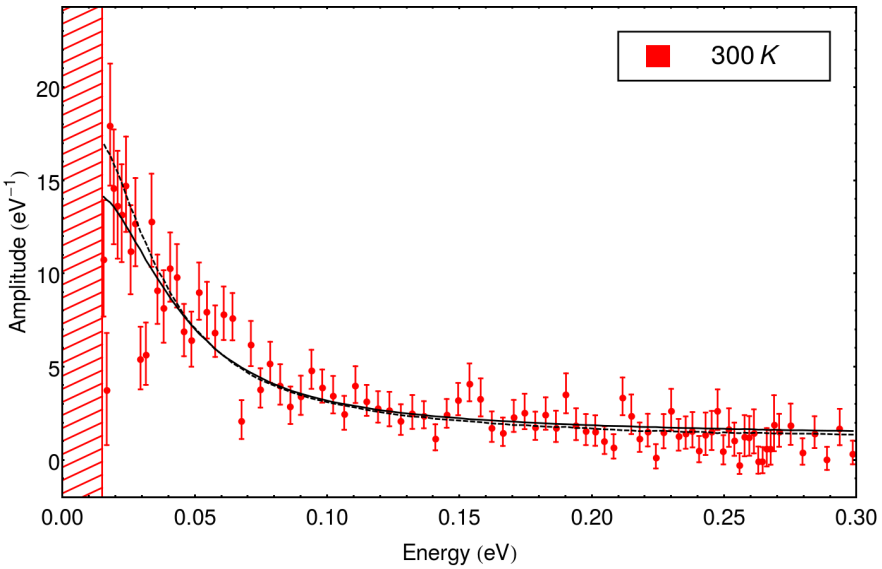


Figure 2.17: Energy spectrum of the **300K** measurement (sample A kept at 300 K bombarded with 7 keV positrons) with superimposed the best fit from out model using $1/p = 100000$ (continuous black) and the temperature-dependent p (dashed black).

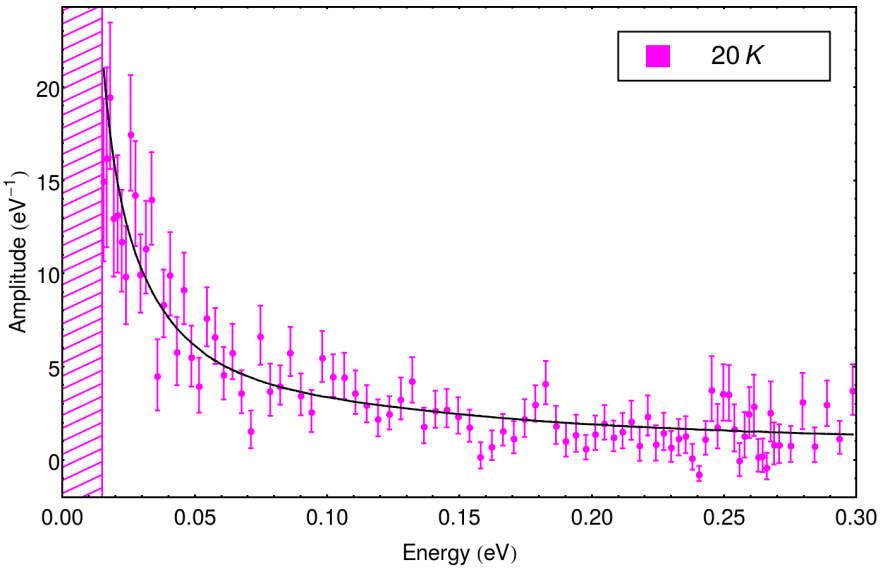


Figure 2.18: Energy spectrum of the **20K** measurement (sample A kept at 20 K bombarded with 7 keV positrons) with superimposed the best fit from out model using $1/p = 100000$ (continuous black) and the temperature-dependent p (dashed black).

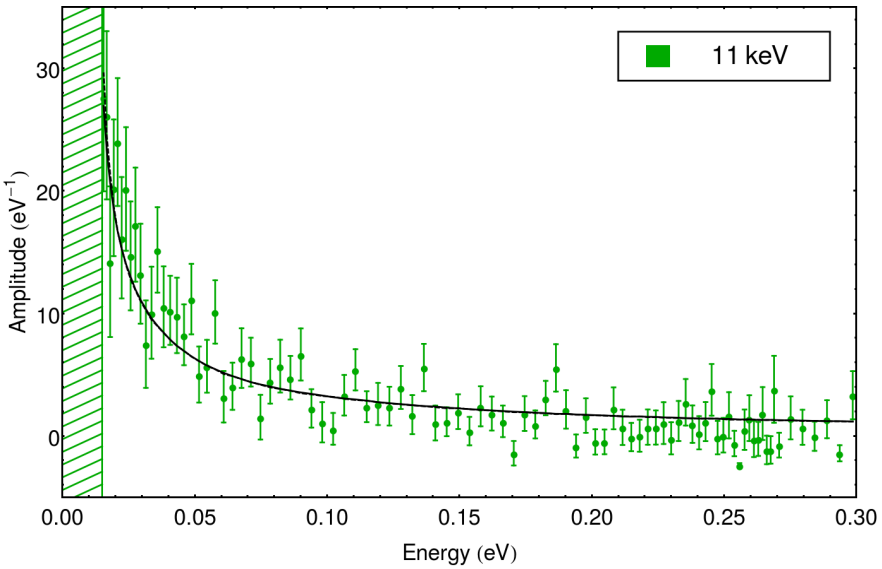


Figure 2.19: Energy spectrum of the **11K** measurement (sample A kept at 20 K bombarded with 11 keV positrons) with superimposed the best fit from out model using $1/p = 100000$ (continuous black) and the temperature-dependent p (dashed black).

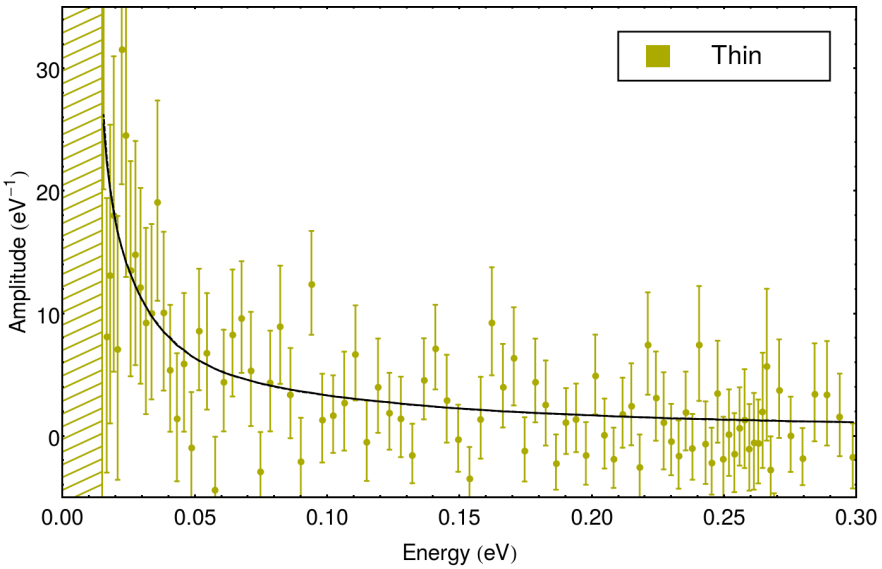


Figure 2.20: Energy spectrum of the **Thin** measurement (sample B kept at 20 K bombarded with 7 keV positrons) with superimposed the best fit from out model using $1/p = 100000$ (continuous black) and the temperature-dependent p (dashed black).

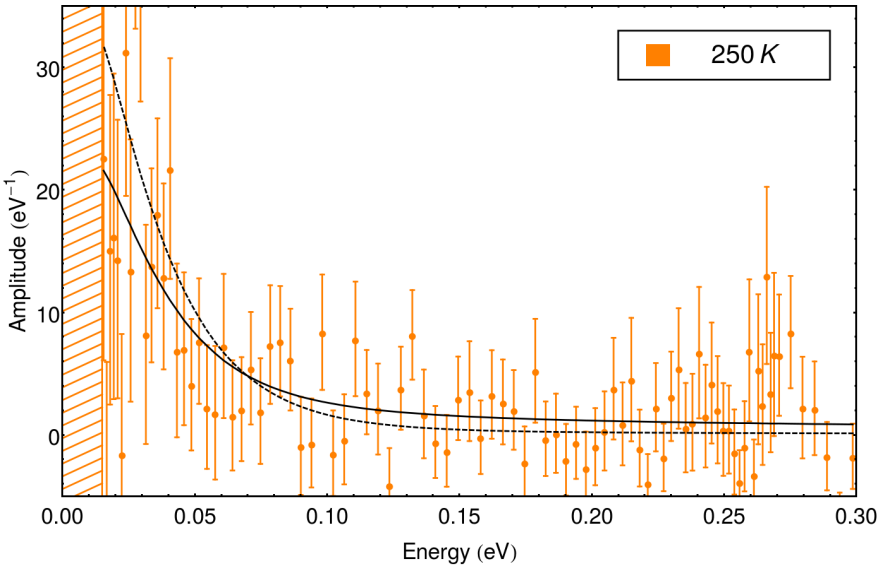
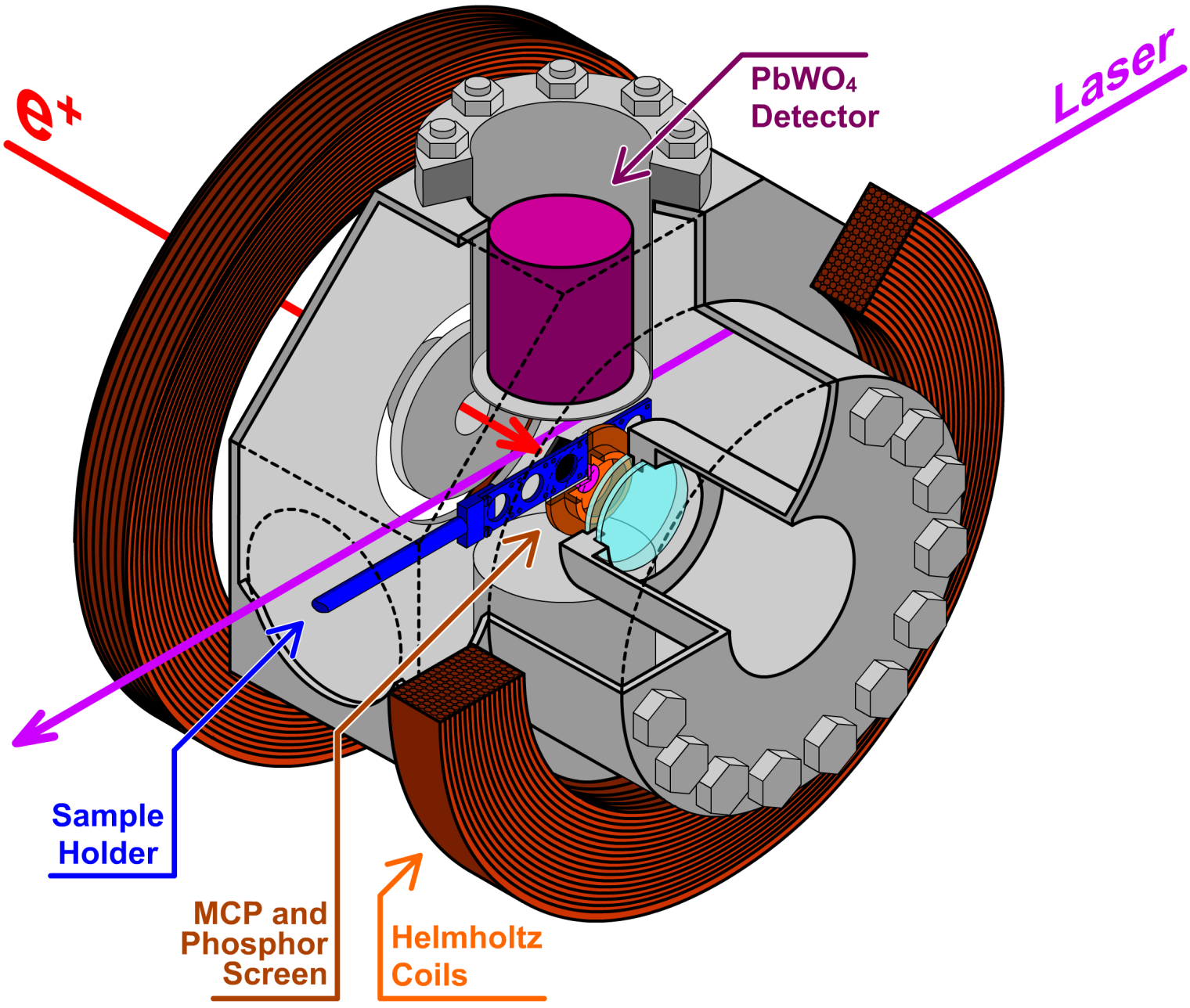


Figure 2.21: Energy spectrum of the **250** measurement (sample B kept at 250 K bombarded with 7 keV positrons) with superimposed the best fit from out model using $1/p = 100000$ (continuous black) and the temperature-dependent p (dashed black).

Sectional view of the secondary positron chamber (also known as the Breadbox*) in which the spectroscopy experiments detailed in this chapter have been carried out.*


e+
PbWO4
Detector
Laser
Sample
Holder
MCP and
Phosphor
Screen
Helmholtz
Coils

Chapter 3

Ps Spectroscopy in AE$\bar{g}$IS

It is black and dark by the privation of light. Doth not the light comfort all the world?

– François Rabelais
Gargantua and Pantagruel

Employing the charge exchange scheme[104] to produce antihydrogen the AE$\bar{g}$IS experiment needs a reliable intense pulsed source of cold positronium, which translates in the need for the apparatus to feature a dedicated positron line[105].

The AE$\bar{g}$IS positron line orignates from a 50 mCi ^{22}Na radioactive source which provides a constant flux of positrons with an endpoint energy of 546 keV[41]. These positrons are then moderated to an energy of a few eV through a solid neon moderator[42] which is grown on a conic shaped support and, shortly after growth, presents an excellent moderation efficiency (exceeding 0.7%). During operation the moderator efficiency will then degrade with a rate that is dependent on the amount and type of residual gas present in the vacuum chamber. A standard operation procedure is to periodically heat, evaporate, cool and regrow the moderator; the procedure goes by the name of *regeneration* and takes around 40 minutes to be performed.

Moderated positrons are captured then inside of a Surko type trap[106] which further cools down the trapped positrons by means of interaction with gaseous N_2 introduced in the trap. The Surko style trap is loaded continuously and unloaded every 152 ms into a subsequent Malmberg-Penning trap that we call the *Accumulator*. The *Accumulator* is held at a pressure of about 10^{-10}mbar which gives us the possibility to store positrons for times in the order of hundreds of seconds. The amount of positrons trapped into the

Accumulator increases with the number of accumulated pulses up until it saturates at around $8 \cdot 10^7$ which during commissioning, when a 11 mCi source was employed, happened after 3000 pulses were accumulated[105] and at the moment happens after 600 accumulated pulses. After the accumulation phase is concluded the positron cloud is then extracted from the accumulator into the transport line. Considered up to this stage the positron line produces bunches of $\approx 8 \cdot 10^7$ positrons with an energy of 300 eV, a length of $15 \div 20$ns and a repetition rate of $\approx 0.011\,\mathrm{s}^{-1}$.

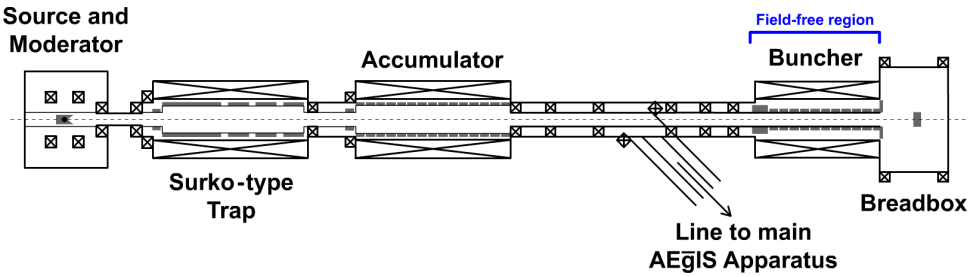


Figure 3.1: Schematic view of the AEgIS positron line. During normal operation the beam is steered through the oblique transfer line into the main trap line, alternatively positrons can be bunched and used in the secondary positron chamber (*Breadbox*).

3.1 The secondary positron chamber

During the nominal AEgIS apparatus operation (that is production and use of $\bar{\mathrm{H}}$) positron bunches extracted from the *Accumulator* are magnetically transported into the main trap line. Indeed our apparatus allows the possibility, whenever positrons are not needed in the main trap line, to instead transport them through another line, compress them with a buncher to 7 ns long bunches, accelerate them to an energy of a few keV and employ them in the secondary positron chamber that we refer to as the *Breadbox*.

The *Breadbox* (depicted in the heading page of this chapter) is an octagonal prism shaped vacuum chamber in which positron bunches are injected along the prism axis. A manipulator-operated sample holder can be loaded with up to 4 samples and allows them to be placed on the focus of the positron beam. When the sample holder is moved away from the positron trajectory, positron bunches will impinge onto an MCP coupled with a phosphor screen. The phosphor screen is visible through a viewport installed onto the back of the *Breadbox* and is shot with a digital camera to obtain an imaging of the positron spot. On opposite sides of the octagonal chamber we installed two cups that protrude into the experimental chamber reaching a distance of 40 mm from the target center. The cups are used to hold scintillating crystals coupled

to photomultiplier (PMT) tubes that are used to detect γ rays originating from positron or positronium annihilations inside of the chamber. The entire chamber is immersed in a uniform magnetic guiding field generated by a pair of Helmholtz coils.

The signal coming from the scintillator detectors is digitized through an oscilloscope and the resulting waveform, along with the images acquired through the digital camera are collected by a dedicated server called *Captor*. The software running on *Captor* has been entirely written specifically to execute the *Breadbox* data acquisition. The server offers a web interface that is used by the operators to control the data acquisition process and to perform online data analysis. A detailed description of the *Captor* tool is given in appendix C.

The first purpose that the *Breadbox* has served has been the commissioning in late 2014 of the positron line[105], during which the pulse width had been characterized with PbF_2 scintillators and the alignment and size of the spot has been measured through the phosphor-screen-coupled MCP captured with a fast CMOS camera[1]. Both measurements led to fine tune optimization of the positron apparatus.

After that the *Breadbox* started serving its own scientific purpose into spectroscopic study of positronium physics. The two sides of the octagon sitting orthogonally to the cup-mounting ones feature viewports through which laser pulses from the AE$\bar{g}$IS laser apparatus can be shone into the chamber.

The AE$\bar{g}$IS apparatus features three lasers: the *UV*, the *Rydberg* and the *Ionization* laser. The *UV* laser has a fixed wavelength of 205 nm, tuned to excite positronium from the 1^3S to the 3^3P state. The *Rydberg* laser has a tunable wavelength that can vary between 1650 nm and 1720 nm; this range has been chosen as to allow the laser to be used to excite Ps from the 3^3P state to Rydberg band states comprised between $n = 14$ and $n = 18$, hence the name. Finally the *Ionization* laser has a fixed wavelength of $\lambda = 1064$ nm which is enough to completely ionize a 3^3P positronium. Different lasers can be shot in quick succession in any combination to produce Rydberg excited positronium or to perform spectroscopy investigations.

3.2 SSPALS spectra

Our main tool to perform spectroscopy experiments on positronium is the Single Shot Positron Annihilation Lifetime Spectroscopy (SSPALS) spectrum as firstly proposed by Cassidy in 2006[107]. SSPALS spectra consist in the measurement of the annihilation rate of positrons and positronium inside of the *Breadbox* chamber as a function of time. During our spectroscopy inquiries

[1] Hamamatsu Flash4 v2

we acquired SSPALS spectra using two $PbWO_4$ scintillator detectors acquired through a digital oscilloscope.

If an aluminum target, which does not produce positronium, is placed in the chamber, the $PbWO_4$ spectrum appears as a single peak having a width comparable to the length of the positron packet followed by a very faint tail, having typically an intensity at least two orders of magnitude lower than the peak and an irregular shape that depends mainly on the chamber geometry (see the blue curve in figure 3.2). This tail encompasses all of the unmanaged behaviors in the positron line (such as delayed positrons and spilling or reflection during transport) and backscattered positrons from the target that annihilate onto the chamber walls.

If, instead, a nanochanneled silica target is placed in the experimental chamber the SSPALS spectrum displays a long lived exponentially shaped tail following the prompt peak with an initial intensity usually one order of magnitude lower than the prompt peak. The tail derives from low temperature positronium atoms being emitted from the target surface annihilating in flight inside of the experimental chamber and, at higher times, against the chamber walls. The annihilation against the chamber walls can manifest (if the signal to noise ratio is high enough to allow its detection) as an excess of signal above the exponential profile.

Examples of SSPALS spectra acquired either with an aluminum target or a silicon NCP target are shown in figure 3.2.

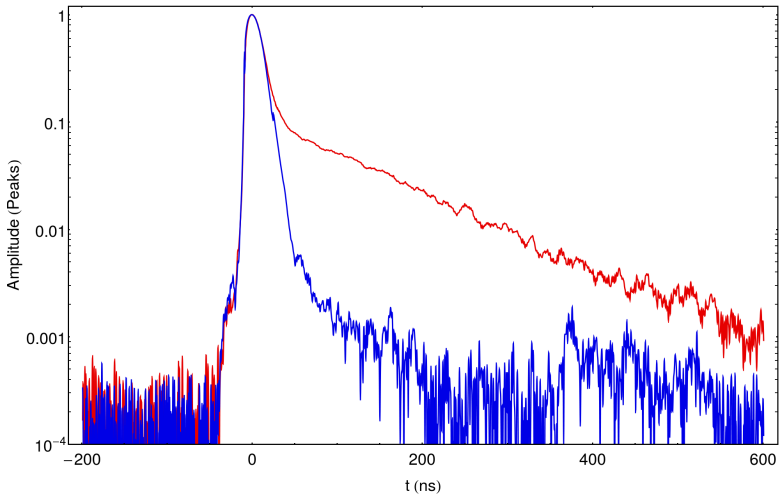


Figure 3.2: Example of SSPALS spectra taken with an aluminum target (blue) and with a silicon NCP (red), the increase in the tail is the positronium signal whose exponential decay appears linear in the semilogarithmic scale. The vertical scale has been normalized to the peak height.

3.3 Producing n = 3 positronium

To circumvent the prohibitive technical difficulties posed by the construction of a laser capable of efficiently exciting positronium from the 1^3S level to the Rydberg band[104], the excitation of Ps in AE$\bar{g}$IS is done in two steps: 1^3S $\rightarrow$ 3^3P $\rightarrow$ Rydberg band. This design is a novelty since in every previous experiment the production of Rydberg positronium had always been performed by transitioning from $n = 2$ states instead.

In 2015 we demonstrated for the first time the excitation of 3^3P positronium levels and then from there the excitation to the Rydberg band[101]. The measurement was performed by implanting positron bunches of $3 \cdot 10^7$ positrons with an energy of 3.3 keV in a nanochanneled silicon converter and measuring SSPALS spectra over the resulting positronium clouds by means of a $20 \times 25 \times 25$mm $PbWO_4$ scintillator placed into the upper detector holder cup coupled to a photomultiplier tube[2] and digitalized through an oscilloscope with a 500 MHz bandwidth[3].

Since SSPALS spectra range in amplitude over several decades, the PMT signal was acquired on different channels with amplifications differing by a factor of 10. The resulting digitized curves capture different parts of the spectrum: the curve acquired with lower gain is able to effectively sample the prompt peak but shows poor precision on the positronium tail; the high gain curve is incapable of recording the prompt peak due to saturation but has excellent precision in sampling the positronium tail. By merging together the two curves (see figure 3.3) the entirety of the spectrum can be recorded effectively.

A series of positronium bunches were shot on the target; on even shots the emerging positronium clouds were exposed to both the *UV* and the *Ionization* laser, whereas no laser was shone onto the positronium clouds of odd shots. All of the resulting SSPALS spectra were normalized to the peak height; then even shots were averaged into a single *Laser On* spectrum and odd shots were averaged into a single *Laser Off* spectrum, which are both shown in figure 3.4. On both average spectra we integrate the area under the curve in the 50 ns $\div$ 250 ns time window (measured with the zero placed at the prompt peak); the range has been chosen so as to not include neither the prompt peak nor latter times at which positronium annihilation onto the chamber walls becomes significant.

Any change in the positronium lifetime induced by laser excitation will reflect onto the area measured under the SSPALS spectrum. Assuming the solid angle acceptance of the scintillating detector depends weakly enough on the position inside of the chamber in which the annihilation takes place (which is a safe assumption in our setup) the SSPALS spectrum is a measure

[2] Hamamatsu R11265-100
[3] Tektronix TDS5054B

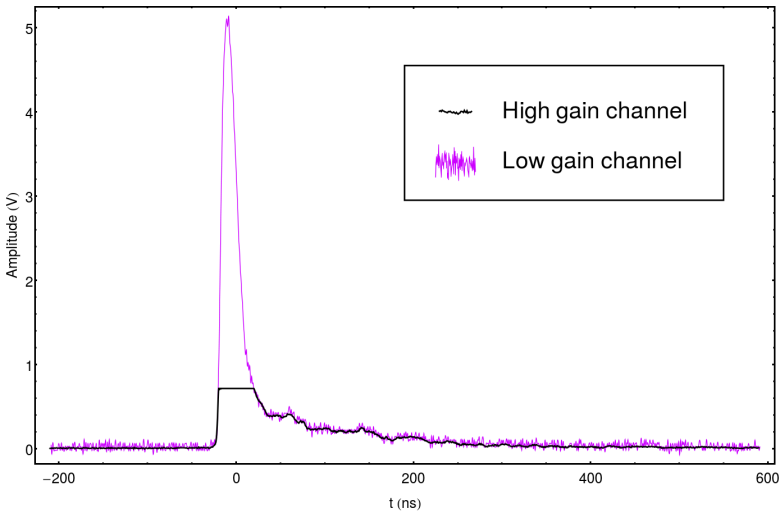


Figure 3.3: An SSPALS spectrum acquired with two different gain levels. The high gain curve lacks the prompt peak measurement due to saturation while the low gain curve does not allow for a precise measurement on the SSPALS tail. In our mesurements we merge together the two curves to harness the capabilities of both gain levels.

of the instantaneous $\mathrm{Ps} + e^+$ annihilation rate.

In the selected integration region the annihilation rate is expected to follow an exponential law:

$$R(t) \;=\; \frac{N}{\tau} \cdot e^{t/\tau} \tag{3.1}$$

Where N is the number of produced Ps and $\tau = 142\,\mathrm{ns}$ is the ortho-positronium lifetime. Since we are operating within the breadbox which features a guiding magnetic field of 25 mT we expect most $n = 3$ positronium to follow two main decay paths: either it will decay by first transitioning back to the $1^3\mathrm{S}$ state or, due to the magnetic field, it would transition to the $1^1\mathrm{S}$ and self-annihilate almost immediately ($\tau_p \approx 125\,\mathrm{ps}$). We expect, if we were to shine only the UV laser, the decay spectrum to become:

$$R^*(t) \;=\; B \cdot \epsilon_P \epsilon_G \cdot \frac{N}{\tau + \tau_f} \cdot e^{t/(\tau+\tau_f)} \;+\; (1 \,-\, \epsilon_P \epsilon_G) \frac{N}{\tau} \cdot e^{t/(\tau)} \tag{3.2}$$

where ϵ_G is the component of the excitation efficiency to the $n = 3$ level due to the geometric factors (namely the limited overlap of the laser beam and the positron cloud) and $\epsilon_P \approx 1/2$ is the component of the excitation

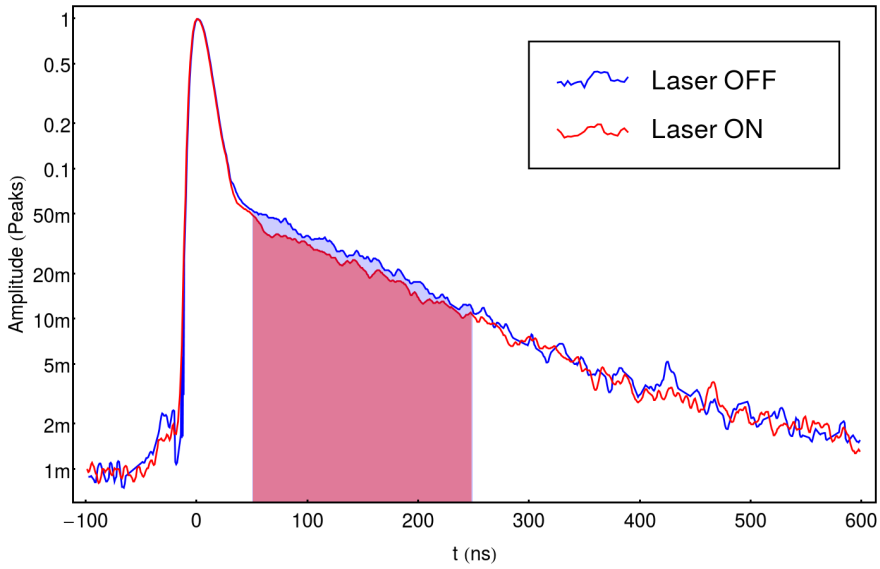


Figure 3.4: SSPALS spectra demonstrating the successful excitation of positronium to the $n = 3$ level. The highlighted region is the time range where the curves have been integrated to produce the $\mathcal{A}_{on}$ and $\mathcal{A}_{off}$ parameters.

efficiency due to the excitation dynamics (in our case mainly the fact that the laser stimulates both the excitation and the de-excitation of the $n = 3$ level). $\tau_f \approx 20\,\text{ns}$ is the optical lifetime of $n = 3$ positronium and B is the branching ratio of $n = 3$ positronium to decay through the 1^3S transition. Even in the (unrealistic) situation in which we were to produce with perfect efficiency the $n = 3$ state ($\epsilon_G \epsilon_P = 0.5$) we would expect a variation in the area under the spectrum of less than 1% which, as we will see in section 3.5 is technically extremely difficult to detect.

Instead if we shoot also the *Ionization* laser the expected spectrum becomes:

$$R^{**}(t) \;=\; B \cdot \epsilon_P^* \epsilon_G \cdot (1 - \epsilon_I) \cdot \frac{N}{\tau + \tau_f} \cdot e^{t/(\tau+\tau_f)} \;+\; (1 - \epsilon_P^* \epsilon_G)\, \frac{N}{\tau} \cdot e^{t/(\tau)} \tag{3.3}$$

where ϵ_I is the ionization efficiency of the $n = 3$ level and ϵ_P^* is one again the component of the efficiency due to the excitation dynamics, modified to account for the depletion of the $n = 3$ level caused by the ionization laser. Considering the power (50 mJ) the spot size and the absence of doppler broadening effects in the ionization process we can safely conclude that most $n = 3$ positronium is ionized by the laser pulse, ($\epsilon_I \approx 1$). Consequence of this and of the fact that $\epsilon_P \approx 1/2$ (resulting from computations by Dr.Caravita[88])

is that $\epsilon_P^* \approx 1$. Equation 3.3 then becomes:

$$R^{**}(t) \;=\; (1 \,-\, \epsilon_G)\frac{N}{\tau} \cdot e^{t/(\tau)} \;=\; (1 - \epsilon_G) \cdot R(t) \tag{3.4}$$

Therefore from $R(t)$ and $R^{**}(t)$ we can compute the geometric component of the excitation efficiency ϵ as:

$$\epsilon_G \;=\; \frac{R(t) \;-\; R^{**}(t)}{R(t)} \tag{3.5}$$

Although on a purely algebraic point of view ϵ_G can be computed from the height of SSPALS spectra at any time t, mitigation of the experimental uncertainty makes it mandatory to instead integrate the SSPALS curve in a suitable region to damp high frequency noise effects. To estimate the excitation efficiency we therefore employ the parameter S defined as:

$$S \;=\; \frac{\mathcal{A}_{on} - \mathcal{A}_{off}}{\mathcal{A}_{off}} \tag{3.6}$$

Where $\mathcal{A}_{on}$ and $\mathcal{A}_{off}$ are, given a time range, the integral of the SSPALS spectrum in that time range normalized to the number of produced Ps atoms. The normalization in this case is given, as said before, by dividing the SSPALS spectrum by the prompt peak height, since the prompt height gives an estimation of the intensity of the positron shot intensity which is precise enough to perform this measurement. As we will see in section 3.5 this is not the case for higher precision measurements.

Our measurement resulted in an S parameter of 0.155 ± 0.011 which indicates the successful production of $n = 3$ positronium with a efficiency of about 15% when the *Ionization* laser is fired. Consequently we can expect, in its absence, to have an $n = 3$ population of about 8% of the entire positronium cloud when only the *UV* laser is fired. Confirmation that this change in the S parameter is effectively due to $n = 3$ positronium excitation can be found by changing the tuning of the UV laser and observing that the S parameter is indeed dependent on the laser frequency. A scan of the S parameter as a function of the UV laser frequency is shown in figure 3.5: data clearly shows the excitation resonance.

3.4 Producing Rydberg positronium

We excite positronium into Rydberg states by shooting synchronously the *UV* and *Rydberg* laser onto a positronium cloud. The Rydberg states we are

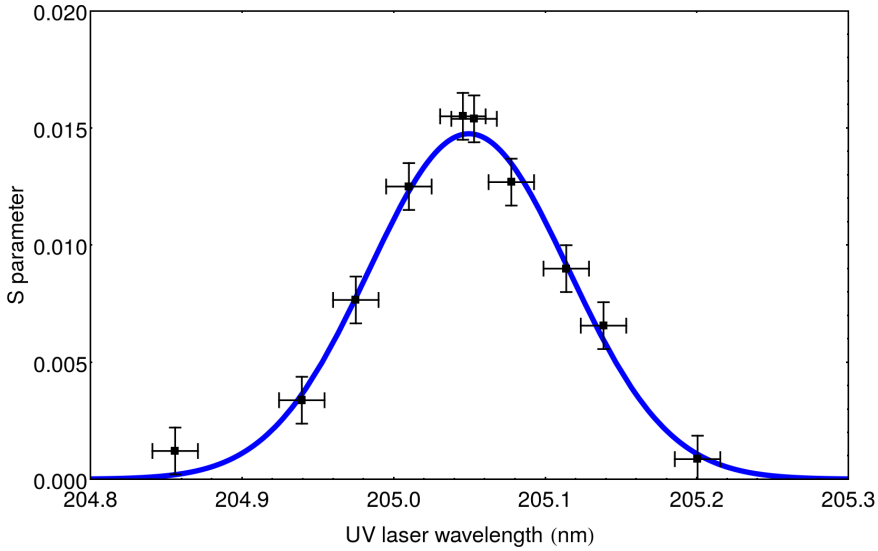


Figure 3.5: Measured value of S during the production of $n = 3$ positronium as a function of the UV laser frequency, displaying the excitation resonance at 205.05 ± 0.02nm

interested in forming ($15 \leq n \leq 18$) have lifetimes in excess of 10 µs[44], which allows positronium produced by the NCP to travel for more than 50 cm if we assume an average speed of $5 \cdot 10^4$m/s for the produced Ps. In our experimental apparatus this means that a large portion of the Rydberg-excited positronium will reach the chamber walls and annihilate onto them, causing an excess on signal at higher times.

The experimental chamber walls closest to the sample are the bottom surfaces of the two detector holder cups. Their distance from the beam spot center ranges from 2 cm to 5.2 cm which means that we expect their pickoff effect to become visible from around 400 ns to around 1 µs. This is confirmed by the experimental data; in figure 3.6 we plotted the average of several SSPALS spectra acquired with no laser being shone onto the positronium cloud, multiplying the curve point by point by a factor $e^{t/\tau}$ with $\tau = 142$ ns. This transformation turns exponential decays into flat lines and makes the excess of signal due to Ps annihilations against the chamber walls clearly visible from 400 ns onwards.

We can employ the excess signal at $t > 400$ ns to detect the successful formation of Rydberg positronium. To do so we measured the S parameter as defined in equation 3.6 using the area under the curve in the 300 ns to 600 ns range and repeated the measurement employing different frequencies for the *Rydberg* excitation laser. The result is shown in figure 3.7: the $n = 15$ and $n = 16$ levels are clearly visible, while levels $n = 17$ and $n = 18$ are just hinted

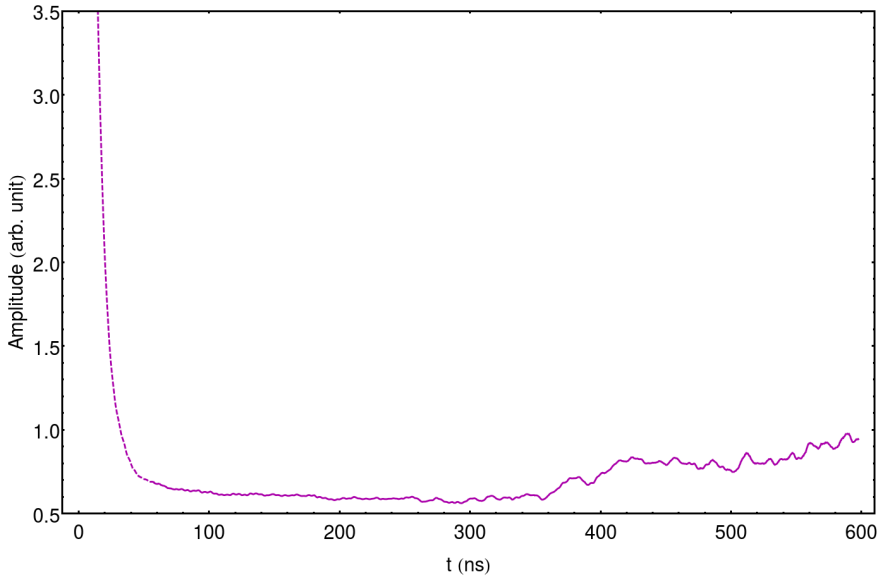


Figure 3.6: SSPALS spectrum normalized through multiplication by $e^{t\cdot\tau}$ with $\tau = 142\,\text{ns}$. The excess above the exponential decay due to pickoff from the chamber walls is clearly visible beyond $t = 400\,\text{ns}$

due to Rydberg band broadening caused both by the magnetic field (25 mT) inside of the *Breadbox* and by the broad band that the laser features by design: it is in fact the main goal of the *Rydberg* laser to excite positronium inside of the production region which is permeated by a 1 T magnetic field that causes the complete overlap of the Rydberg bands, under this circumstance a broad band laser is capable of a more efficient excitation.

3.5 Detecting n=2 positronium

We anticipated in section 3.3 that most 3^3P positronium annihilates by first transitioning to the 1^3S or 1^1S state. The main competing decay process is the $3^3\text{P} \rightarrow 2^3\text{S}$ transition through spontaneous emission. We expect the latter decay process to have a branching ratio of 12% [88, 108] (see figure 3.8 for a scheme of the relevant positronium excitation levels). The 2^3S positronium state is expected to have a lifetime of 1.14 µs[109] therefore it should appear in the SSPALS spectrum in the form of a long-lived component; we'll employ again the tools described in the previous paragraphs and the S parameter to attempt to detect metastable 2^3S positronium: if successful this represents a novel path to producing 2^3S positronium as opposed to the alternative laser excitation of the 1^3S level in presence of external electric fields[110].

Beyond its intrinsic scientific interest, the production of this excitation is

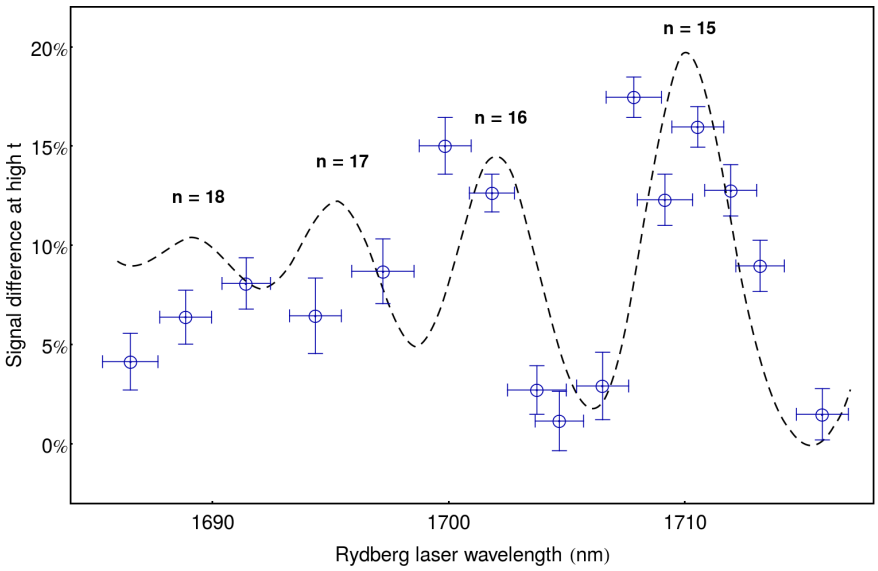


Figure 3.7: S parameter computed in the $300 \div 600$ ns range as a function of the *Rydberg* laser wavelength, demonstrating the successful excitation of Ps to the Rydberg band.

of great relevance since $n = 2$ positronium has been proposed as a probe to test the free fall of mixed matter-antimatter systems [111, 112, 113, 114] or to perform fundamental QED tests [115, 116, 117].

The measurement technique consists in choosing two regions of interest in the SSPALS spectrum timeline; the first must include the times at which we expect the positronium to be flying in the chamber, avoiding the prompt peak and the latter positronium-walls interactions, the second region must be chosen to include the annihilations happening when positronium reaches the chamber walls. We found suitable choices for the two regions to be $70 \div 350$ ns from the peak for the first area and $350 \div 500$ ns for the second. The experimental technique we employ is to perform two measurements: one alternating absence of laser light and the shining of the UV laser, the other alternating absence of laser and the shining of both the UV and the *Ionization* laser; we then determine the S parameter in both ranges for both measurements. If no $n = 2$ metastable state is produced we expect (in both measurements) the S parameter to be the same in both regions since in both cases the laser light would be removing some positronium but not changing substantially its lifetime. On the contrary, if shining solely the UV laser does indeed produce long-lived positronium states, we expect in the UV measurement to observe a lower S parameter in the second region than in the first while still seeing compatible values in the other measurement.

The main difficulty that we faced in performing this measurement is

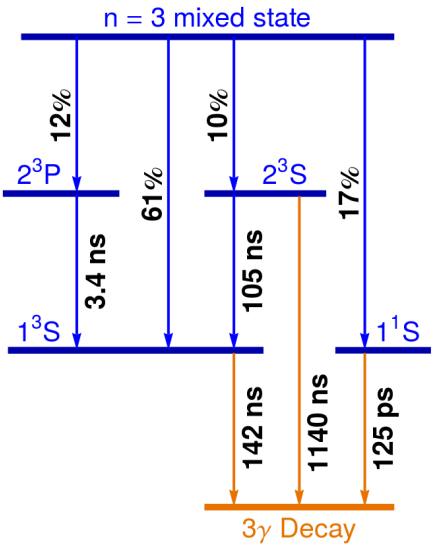


Figure 3.8: Positronium excitation levels relevant for the $n = 2$ excitation measurement. Due to the electric field caused by the last electrostatic lens of the buncher the $n = 3$ sublevels are mixed.

the required precision. Due to the low branching ratio for 2^3S positronium production we expect S parameter variations to be shy of 0.01. To achieve the necessary precision we phased out the oscilloscope employed for the $n = 3$ and Rydberg measurements that featured 8-bit sampling in favor of a 12-bit replacement[4]. Additional RF shielding has been installed onto the apparatus connections and the spent ^{22}Na source (which had an activity of about 11 mCi) was replaced with a new one with an activity of 50 mCi. When acquiring data we increased the statistics from 15 shots averaged for each measurement to over 160. We also introduced a second $PbWO_4$ detector placing it in the lower detector holder, unfortunately data acquired through the second detector proved to have a low signal to noise ratio, therefore we opted to discard it during data analysis. These measures solve all but one problem we will be facing during measurement, that is to correctly normalize the areas to the shot intensity.

As we said earlier in section 3.3 normalization of SSPALS spectra should be performed by dividing them by the shot intensity, i.e. the number of positrons in each shot; the main systematic to be counteracted by this normalization being the aging of the moderator that causes the beam intensity to progressively decrease in time (see figure 3.9). In the previous measurements we employed the height of the prompt peak to normalize SSPALS spectra, the reason for this being that we expect, under similar conditions, every part of the SSPALS

[4]Lecroy HDO6104

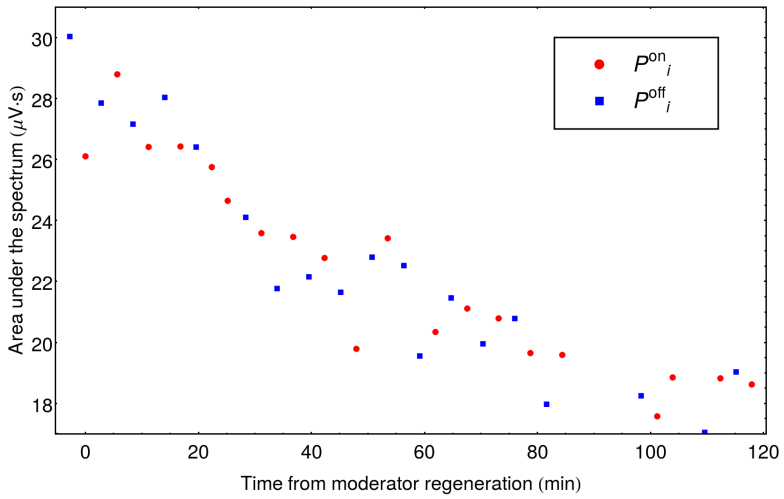


Figure 3.9: Area under the region of the SSPALS curve in the region 370 ns, 650 ns for subsequent shots acquired with alternating laser settings. Blue dots represent measurements that had no laser being shone in the chamber, while the *UV* laser was shone during the acquisition of the measures plotted in red. The x axis shows the time elapsed since the last moderator regeneration: the decrease in the beam intensity due to the moderator aging is patent.

spectrum to be proportional to the implanted shot intensity and the peak height is a precisely determinable feature that is not affected by the employed lasers. We observed that due to the increased shot statistics our detector response at high intensities ceases to be linear. This nonlinearity can be seen clearly when areas under tails of SSPALS spectra are compared among each other and with the peak as shown in figure 3.10.

A traditional way to solve this difficulty could be to attempt an interpolation of the nonlinear behavior shown in figure 3.10 to compute a correction factor for the peak normalization. If this route is followed the final result will suffer in the first place by the uncertainty introduced by the approximated normalization correction factor, in the second place by the unavoidable uncertainties introduced by merging curves sample with different gain; moreover will present no clear way of determining the experimental uncertainty on the measured S parameter. Instead we opted for developing a completely novel normalization approach[118] that will make use only of the portion of the curve acquired at high gain and will offer a natural way of determining the final uncertainty on the S parameter. We named this technique *Detrending* due to its similarity with the homonymous frequency spectrum analysis technique.

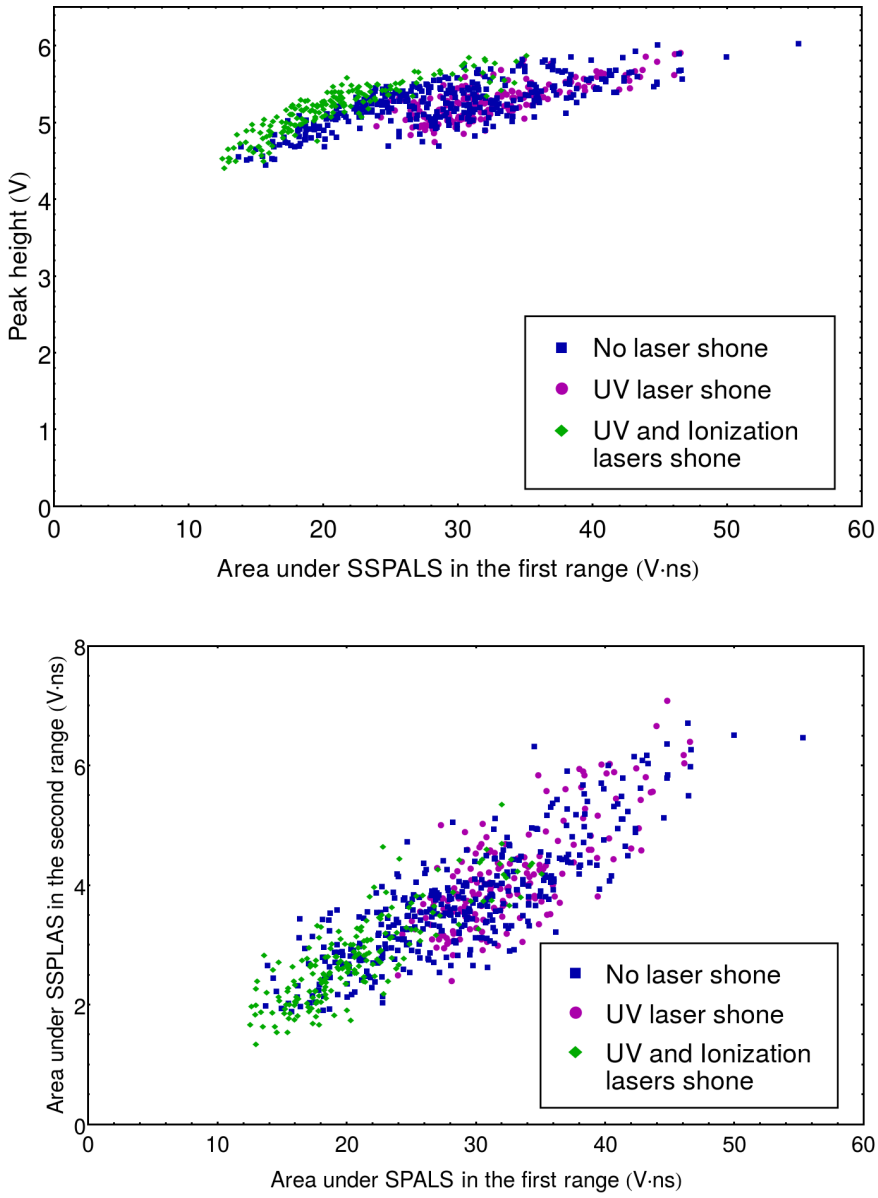


Figure 3.10: Comparison between the peak height and the area under the SSPALS curve in the first region (top) and between area under the SSPALS curve in the two selected regions (bottom). Colors encode for different laser conditions. In the comparison of peak against area the non linearity of the peak is clearly evident, whereas area measurements in the tail appear to be linear although they differ in the input intensity by almost a decade.

3.5.1 Detrending

We will here describe the detrending technique in its most general formulation; a more concise exposition can be found in[118].

Let's consider a parameter $P(I, b)$ (in our case the area under a region of the SSPALS curve) that depends upon other two parameters I and b where I is a real number (in our case the shot intensity) and b is a boolean condition (in our case the laser status). Let's assume that I is slowly drifting in time (henceforth $I(t)$), therefore making P implicitly dependent on time. Let's also assume that P depends linearly from $I(t)$, which means:

$$\begin{aligned} P(I(t), b) &= k^{on} \cdot I(t) \qquad \text{for} \qquad b = \text{on} \\ P(I(t), b) &= k^{off} \cdot I(t) \qquad \text{for} \qquad b = \text{off} \end{aligned} \tag{3.7}$$

We want to be able to determine the ratio:

$$R = \frac{k_{on}}{k_{off}} \tag{3.8}$$

that is the influence that the boolean parameter b has on the value of P. Let's assume we have measured the value of P at different times with the parameter $b = \text{on}$, obtaining a series of points $\{t_i^{on}, P_i^{on}\}$ where t_i is the time at which the ith measurement with $b = \text{on}$ was performed and P_i^{on} is the measurement result. Similarly let's assume that we have measured the parameter with $b = \text{off}$ obtaining the points $\{t_i^{off}, P_i^{off}\}$.

We start by giving an estimation of the drifting in time of $I(t)$ by approximating a function $f(t)$ proportional to the intensity $I(t)$:

$$f(t) = \alpha \cdot I(t) \tag{3.9}$$

since we are interested in the ratio k_{on}/k_{off} rather than in the absolute value of k_{on} and k_{off} we have no requirement on fixing nor determining the value α but for $\alpha \neq 0$.

We need to first choose a suitable empiric model to describe the function $I(t)$, in our case we saw that a second degree polynomial law suits the task. This model must be used to fit the two series of points $\{t_i^{on}, P_i^{on}\}$ and $\{t_i^{off}, P_i^{off}\}$; let's call the functions resulting from the fit $f^{on}(t)$ and $f^{off}(t)$. We'll employ the arithmetic mean of the two functions as an estimator of the function defined in equation 3.9:

$$f(t) = \frac{f^{on}(t) + f^{off}(t)}{2} \tag{3.10}$$

if the number of measurements performed with $b =$ on differs greatly from the number of measurements performed $b =$ off the mean in equation 3.10 must be weighted according to the accuracy of the two fit results. In figure 3.11 we show the values of $f^{on}(t)$, $f^{on}(t)$ and $f(t)$ for the same series of acquisitions shown previously in figure 3.9.

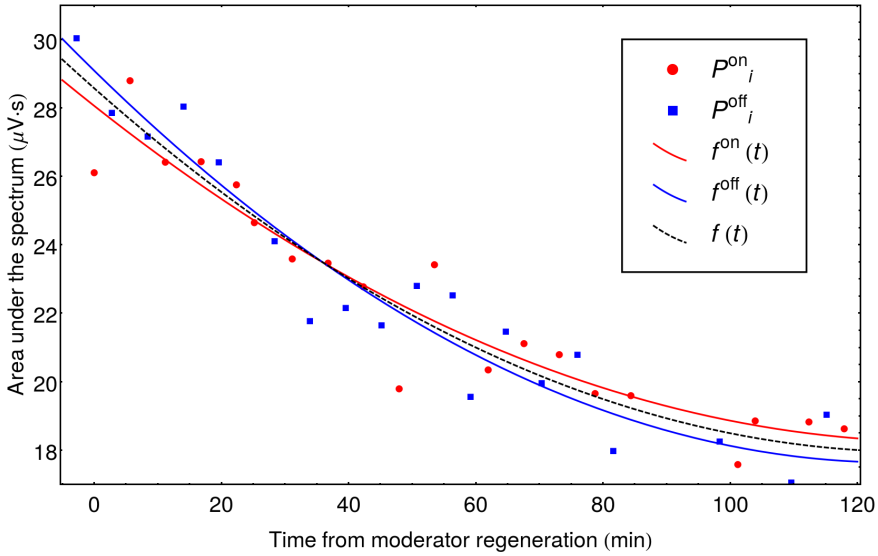


Figure 3.11: We show an instance of the fits to two data series and their mean $f(t)$ which in the detrending procedure is employed as an estimation of the moderator aging.

We then divide the values of $P_i^{on/off}$ by the normalization factor $f(t)$ evaluated at the time the measurement was performed. For uniformity with the rest of the chapter and with reference[118] let's call the resulting values $\mathcal{A}_{on}$ and $\mathcal{A}_{off}$, although the validity of the method holds whether the parameter P represents or not an area.

$$\mathcal{A}_i^{on/off} = \frac{P_i^{on/off}}{f(t_i^{on/off})} \tag{3.11}$$

If we use equation 3.7 to substitute P in equation 3.11 we obtain:

$$\mathcal{A}_i^{on/off} \approx \frac{k^{on/off} \cdot I(t_i^{on/off})}{f(t_i^{on/off})} \tag{3.12}$$

where the approximation comes from $P_i^{on/off}$ being a measurement with a finite uncertainty. We then employ equation 3.9 to substitute $f(t)$:

$$\mathcal{A}_i^{\,on/\,off} \;\approx\; \frac{k^{\,on/\,off}\cdot I(t_i^{\,on/\,off})}{\alpha\cdot I(t_i^{\,on/\,off})} \;=\; \frac{k^{\,on/\,off}}{\alpha} \tag{3.13}$$

which means that $\mathcal{A}_{\,on/\,off}$ is independent of time but for the statistic uncertainty in the measurement of $P_i^{\,on/off}$ and the systematic uncertainty in the determination of $f(t)$. Both these uncertainties can be damped by averaging the values of $\mathcal{A}_i^{\,on/\,off}$, let's therefore define:

$$\mu^{\,on/\,off} \;=\; \langle\; \mathcal{A}_i^{\,on/\,off} \;\rangle \;\approx\; \frac{k^{\,on/\,off}}{\alpha} \tag{3.14}$$

The uncertainty $\sigma^{\,on/\,off}$ over the $\mu^{\,on/\,off}$ parameters can be computed as:

$$\sigma_{on/\,off} \;=\; \sqrt{\frac{1}{N(N\,-\,N_{par})}\cdot\left\langle\left(\frac{\mathcal{A}_i^{\,on/\,off}}{f\left(t_i^{\,on/\,off}\right)}-\mu_{on/\,off}\right)^2\right\rangle} \tag{3.15}$$

where N is the number of laser-on (or laser-off) shots and N_{par} is the number of degrees of freedom in the fit performed to compute $f(t)$.

In both equations 3.14 and 3.15 the angled brackets represent an arithmetic mean that should be weighted according to an error model if this is available. To explain how, let's consider for example the case in which an experiment is being performed with a low positron count and very low electric noise. In this case the absolute uncertainty $\Delta\mathcal{A}_i^{\,on/\,off}$ of each measurement is dominated by the shot noise, therefore we can state that, in first approximation, $\Delta\mathcal{A}_i^{\,on/\,off} \propto \sqrt{\mathcal{A}^{\,on/\,off}}$. This value must should be employed, in the examined case, to weight the means employed in the detrending process.

In the case of our experiment the uncertainty appears to be dominated by the electronic noise, therefore we employed a constant error model which reduces the means in equations 3.14 and 3.15 to simple arithmetic means.

The ratio of the μ parameters gives the desired result:

$$\frac{\mu^{\,on}}{\mu^{\,off}} \;\approx\; \frac{k^{\,on}}{\alpha}\cdot\frac{\alpha}{k^{\,off}} \;=\; \frac{k^{\,on}}{k^{\,off}} \;=\; R \tag{3.16}$$

The uncertainty over R can be written from $\mu^{\,on/\,off}$ and their uncertainty over $\sigma^{\,on/\,off}$:

$$\Delta_R \;=\; \sqrt{\frac{\sigma_{on}^2}{N_{on}\mathcal{A}_{off}^2}+\frac{\mathcal{A}_{on}^2\cdot\sigma_{off}^2}{N_{off}\mathcal{A}_{off}^4}}\,. \tag{3.17}$$

In our application we can apply the definition given in equation 3.6 to compute the value of S directly from R as:

$$S = R - 1 \tag{3.18}$$

The absolute uncertainty over S will be the same of R:

$$\Delta_S = \Delta_R - 1 \tag{3.19}$$

3.5.2 Results

We performed the experiment in 2017[118] under non optimal conditions. Due to the impossibility to turn off the last electrostatic lens of the transfer line we experienced a residual electric field in excess of 300 V/cm inside of the *Breadbox* chamber; this influences the energy levels of the Ps atom forcing us to work with mixed states instead of the usual field-free eigenstates, nonetheless the problem has been solved now with the addition of a specialized pulser and the experiment will be repeated in 2018 in an electric field free configuration.

In the presence of an external electric field, due to Stark effect the 2^3S state and 2^3P state mix, considerably shortening the lifetime of the produced $n = 2$ states, which in turn translates into a decrease in the expected signal level. Nonetheless the $n = 2$ states that we form in the electric field present a longer lifetime than the 1^3S and $n = 3$ levels, therefore they are detectable with the proposed technique.

We shot a total of 179 positron bunches while shining the UV laser and a total 161 bunches while shining both the UV and the *Ionization* laser; for each shot in which laser light was shone another one with no laser light was acquired, for a total of 680 acquired shots. Laser on and laser off shots were alternated. During the course of the acquisition of the UV laser shots the moderator has been regenerated 11 times, during the UV + *Ionization* series the moderator has been regenerated 8 times. Each moderator regeneration gives rise to a new data series; different data series are detrended separately and the resulting $\mathcal{A}^{on/off}$ are then averaged all together in a single passage to produce the final result. In figure 3.12 the populations of $\mathcal{A}^{on/off}$ are shown alongside the normal distribution having the same mean and variance of the populations.

We found that when the UV laser was shone alongside the *Ionization* laser the S parameter computed in the first region (0.138 ± 0.006) was compatible with the value computed in the second region (0.149 ± 0.014). Instead when only the UV laser was shone, the S value computed in the first region (0.022 ± 0.005) was incompatible with the value computed at the second region (0.000 ± 0.013). If, in accordance with the model presented in section IV of

[118] we assume as competing models a difference in the S parameter of 0 in the case of absence of signal and of 0.020 in the case of long-lived $n = 2$ mixed stated we can compute the likelihood ratio[119] among the models as:

$$1 \; - \; \frac{\int_{-\infty}^{-0.022} \exp\left[-\frac{(x-\mu^-)^2}{2\sigma^2}\right] dx}{\int_{-\infty}^{-0.022} \exp\left[-\frac{(x-\mu^+)^2}{2\sigma^2}\right] dx} \; \approx \; 90\% \tag{3.20}$$

where $\mu^+ = 0.020$, $\mu^- = 0$ and $\sigma = \sqrt{0.005^2 + 0.013^2} \approx 0.014$

3.6 Applications of detrending

Since its development for the analysis of the data coming from the $n = 2$ excitation experiment, the detrending technique has been increasingly used in the AE$\bar{\text{g}}$IS collaboration to analyze spectroscopic data from positronium experiments and has become de facto the new standard. Due to its success it was also included in the toolset available to perform online positronium excitation data analysis (see appendix C for more details); due to its capability of estimating not only the S parameter but also its uncertainty it is standard practice for the operators to detrend data as it is being acquired, which allows them to decide dynamically the number of shots required to state whether an excitation signal is present or not or if its intensity can be stated with enough accuracy.

During 2017 the *Breadbox* data acquisition and online data analysis server *Captor* has been connected to the plastic scintillating detectors surrounding the 1 T magnet which contains the production trap. This allowed the operators to employ the tools developed for the *Breadbox* measurements to, instead, transport the positron beam onto the conversion target, and align the laser timing and position to excite positronium to the Rydberg band before it diffuses into the production trap.

During the fall of 2017 we were able to confirm the successful excitation to the Rydberg band of the positronium produced in the 1 T magnet. Due to the limited time response of the plastic detectors measuring a high resolution SSPALS spectrum is not possible inside of the 1 T magnet and the long tail of the detector's response makes the Rydberg excitation signal much more faint compared to the background than that measured in paragraph 3.4. Moreover, due to the extreme broadening caused by the 1 T magnetic field, differential measurements such as the one presented in figure 3.7 are expected to result in continuous lines and will not allow to detect the presence of the Rydberg resonances.

Figure 3.13 shows the response of a plastic scintillator (measured employing cosmic rays) with superimposed the signal measured by the plastic scintillators

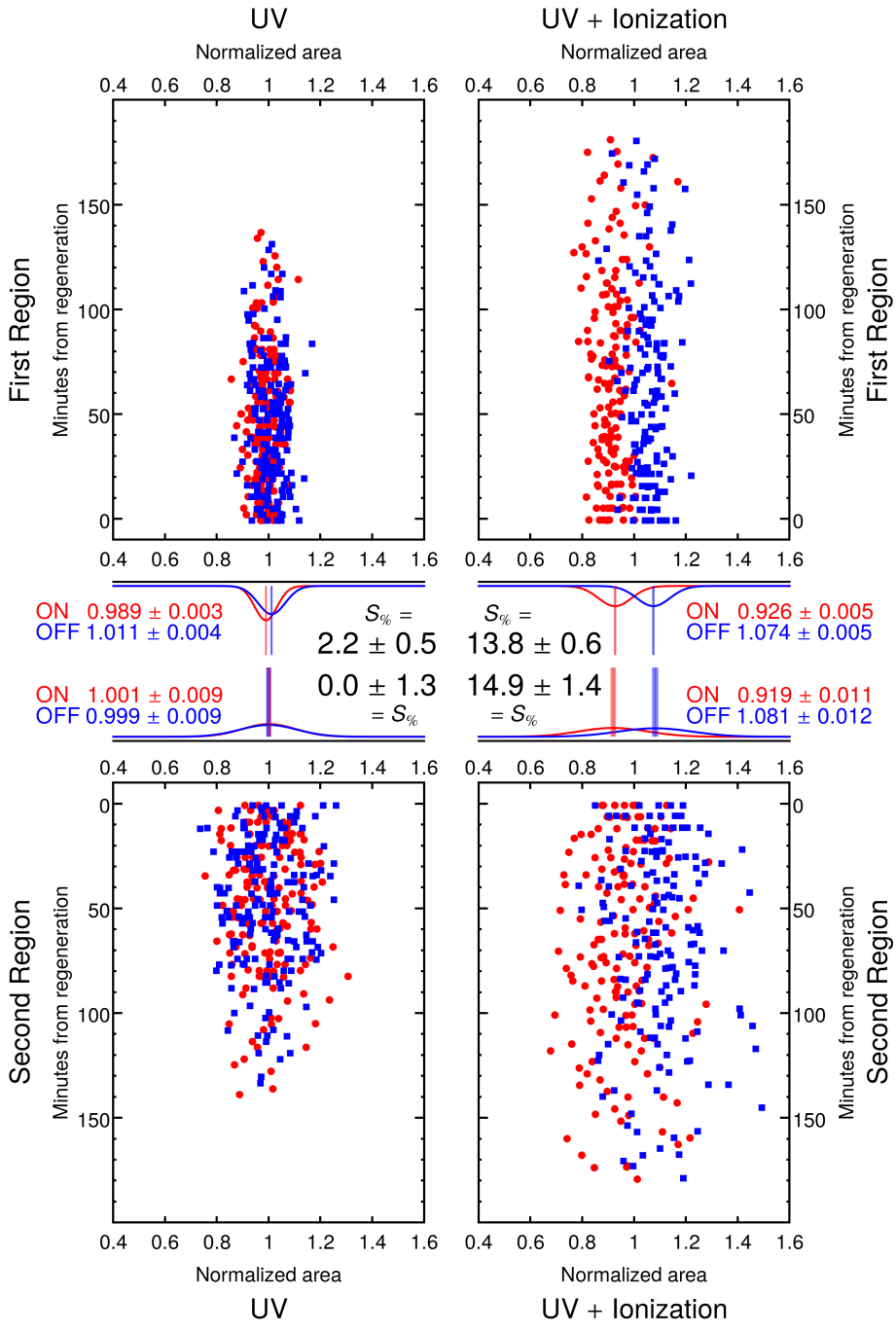


Figure 3.12: Recapitulation of the results of the $n = 2$ measurement. In each frame $\mathcal{A}_{on}$ measurements (red circles) and $\mathcal{A}_{off}$ measurements (blue squares) for the two measurement regions and the two laser settings.

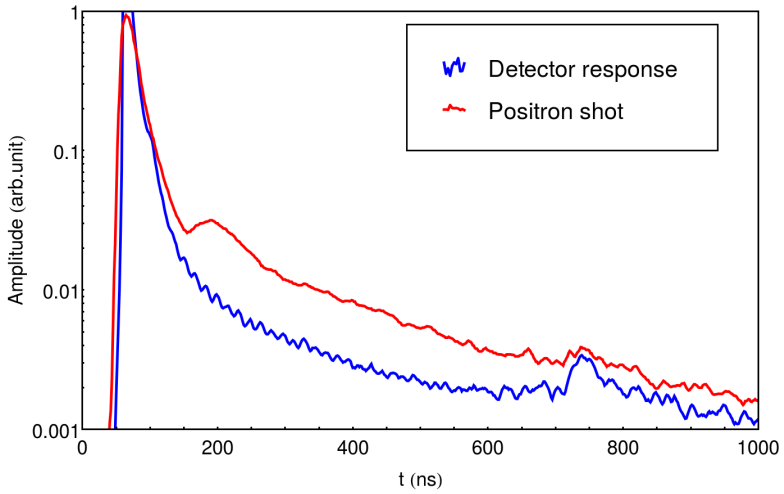


Figure 3.13: In blue the response of the plastic scintillating detectors that wrap around the AEgIS 1 T magnet, in red the signal recorded during positron shots. The excess in signal is due to positronium production by the target. Due to the evident difference in the shape of the prompt peak it is not possible to employ it to normalize these spectra.

during a positron shot with both curves normalized so that their tails at long times ($3\,\mu s < t < 15\,\mu s$) coincide. The excess signal due to the positronium production is clearly visible in the positron shot. Figure 3.14 shows the value of the S parameter computed via the detrending technique on the area under SSPALS spectra integrated over windows of 150 ns. The procedure has been repeated several times moving the integration window in time by steps of 25 ns; the x axis in figure 3.14 displays the position of the center of the window. The excess of signal at high times is due to of long-lived positronium which, similarly to what has been seen in paragraph 3.4, annihilates on the chamber walls rather than in-flight and is evidence of the successful excitation of positronium in the Rydberg band. To get an idea of the geometry of the production trap, we refer the reader to the opening page of chapter 4.

We can confidently say that, without the detrending technique, detecting Rydberg-excited positronium in the production trap would have been near to impossible if not impossible altogether.

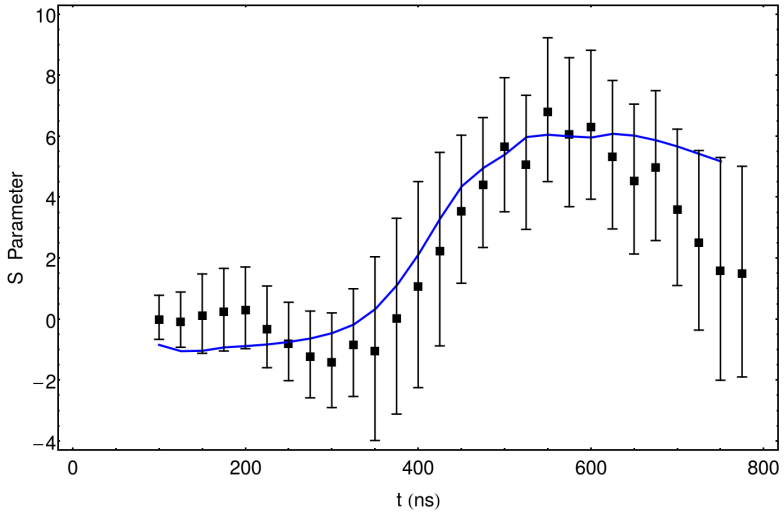
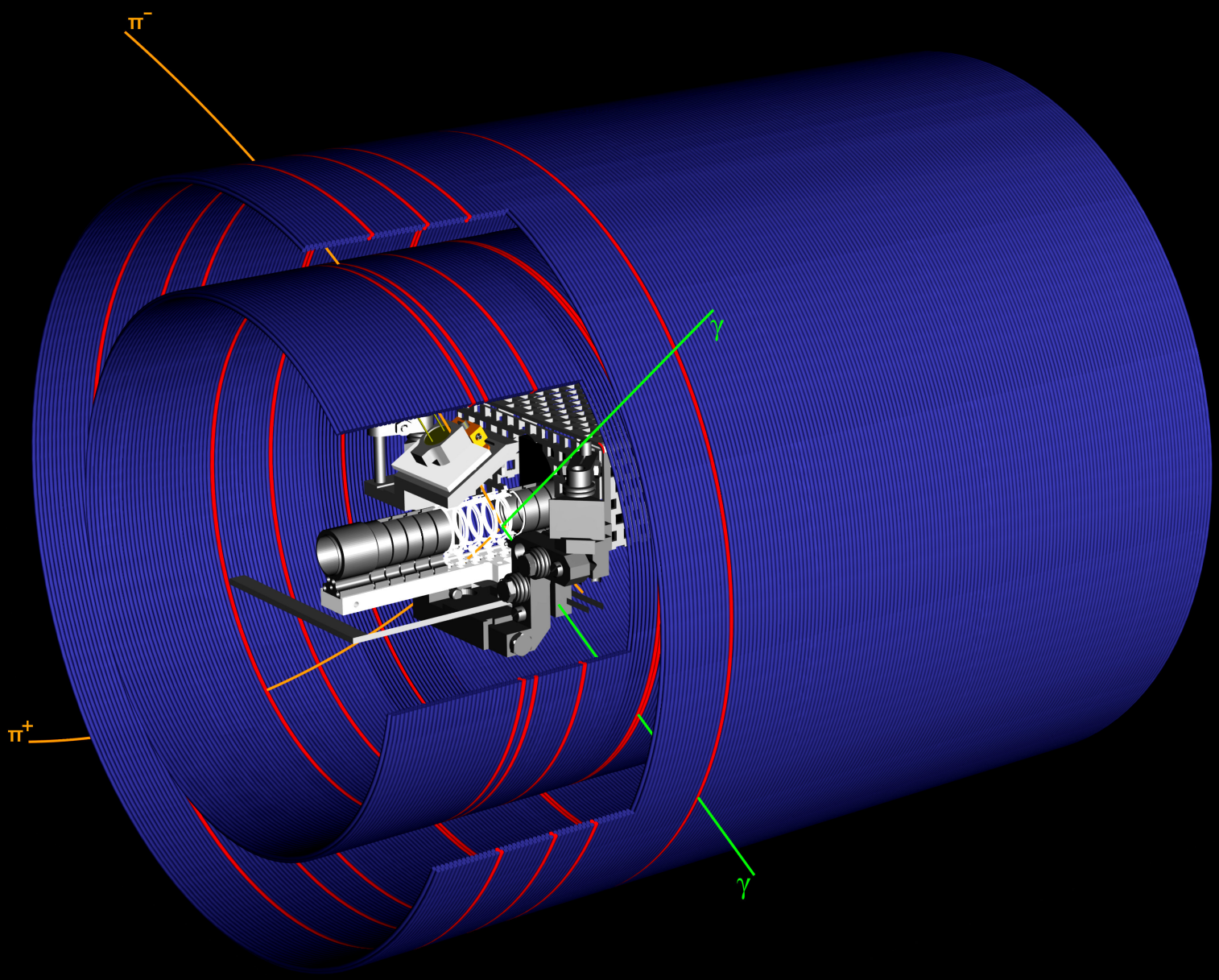


Figure 3.14: In black the S parameter computed for a different 150 ns-wide integration windows through the detrending technique; on the x axis the temporal position of the middle of the window. The blue curve shows the expected value of S computed through Monte Carlo simulations in the case of successful Rydberg excitation (in case of absence of excitation the expected S shows not the increase at 400 ns

The production trap (in gray), located within the FACT detector (in blue). Above the production trap the tilted sample holder that supports the NCP that provides positronium for the charge-exchange reaction. Five electrodes have been depicted in wireframe and a $\bar{p}$ example annihilation on the trap wall is shown. The perfect detection of both $\pi^{\pm}$ and γs by the fibers is an artistic license.


π^-
γ
π^+
γ

Chapter 4

Seeing Antihydrogen

"You never talk of likelihoods on Arrakis. You speak only of possibilities."

– Frank Herbert
Dune

In 2016 the AE$\bar{g}$IS apparatus demonstrated the capability of reliably moderate, capture, cool, compress and transfer $\bar{p}$ from the AD ring into the production trap. The procedure yields a cloud of $4 \cdot 10^5$ $\bar{p}$, a radius (FWHM) of 0.5 mm, a temperature of ≈ 150 K and a lifetime in excess of 1000 s[120]. The positron line was commissioned in 2015 and due to an upgraded transfer line installed in early 2017 is now capable of injecting positron bunches directly onto an NCP target placed above the production trap[1]. In the current working conditions the implantation energy of positrons into the production trap's NCP is non optimal ($E \approx 5$ keV), this results in the production of positronium at ≈ 1000 K. The successful excitation of Ps into the Rydberg band has been also demonstrated inside of the production trap (see section 3.6).

In this condition when the excited Ps cloud crosses the stored $\bar{p}$ cloud we expect the charge exchange reaction to take place and antihydrogen to be produced. Given our current understanding, it is possible that as of 2017 $\bar{H}$ has already been produced inside of the AE$\bar{g}$IS apparatus; at the same time it is not possible to claim it with confidence unless direct evidence of its presence inside of the production trap is shown.

The annihilation of an antihydrogen on a chamber wall or onto residual gas consists in the annihilation of its nucleus and of its positron, therefore the

[1]The original design expected them to be captured again in the main trap line and being further manipulated before undergoing the final implantation into the NCP target.

signature of this process will consist of the emission of the particles that follow the annihilation of an antiproton in coincidence with two gammas from the annihilation of the positron. It is easily seen how an antiproton annihilation can easily mimic an antihydrogen annihilation

The annihilation of an antiproton at rest on a proton or a neutron is accompanied in most cases by the emission of pions (the aggregate pionic decays encompass more than 97% of the branching ratio [121]), with an energy spectrum that peaks around 200 MeV [122]. Pions then decay either into high energy γ rays (π^0) or into muons ($\pi^{\pm}$). Charged pions with energies above 1 MeV have characteristic flight distances before annhilation above 1 m. Therefore as far as the scale of the FACT detector is concerned [2] muon decays are not relevant since, the produced charged pions annihilates inside of the detector. Neutral pions, on the contrary, will decay with a characteristic flight time of a few nanometers, therefore what we might in principle hope to measure with our detector are the gamma rays produced by their annihilation. Unfortunately the efficiency of the FACT fibers in detecting gamma rays ranges in the order of magnitude of 10^{-4}; we can therefore say that we are, in substance, only sensitive to detecting charged pions.

Since in the current conditions we expect to produce an amount of antihydrogen atoms in the order of a few units per run while at the same time the production trap is loaded with a wealth of $\bar{\mathrm{p}}$ we need to address the problem of distinguishing effectively enough between the annihilation of the two species. The difficulty introduced by the fact that the only fundamental difference between an $\bar{\mathrm{H}}$ and a $\bar{\mathrm{p}}$ annihilation (namely the emission of two additional gamma rays) involves particles to which our detector is substantially insensitive to cannot be understated but it is by no means a guaranteed checkmate; detection of antihydrogen employing only charged pions has indeed been proposed before [123]. We will consider the physics of the rest of the experiment and the nature of the operations performed in the trap to perform the species distinction.

4.1 Time-based species discrimination

The first tool that we can invoke to determine the species of annihilating particles in the production trap is the timing of the annihilations.The technique requires only the measurement of the rate of γ production inside of the AE$\bar{\mathrm{g}}$IS apparatus which can be measured using scintillating detectors. The AE$\bar{\mathrm{g}}$IS apparatus features 12 enormous 2 cm thick, 10 cm long slates of polystyrene coiling around the 1 T magnet each revolving for about 180°. These slabs act as scintillating material and are read through PMT tubes and can be used to

[2] The external radius of the detector is less than 10 cm, see nonetheless section 4.2 for a detailed description.

determine the γ rate from the production trap.

To understand the technique let's first illustrate the procedure by which $\bar{H}$ is produced inside of AE$\bar{g}$IS. First of all the production trap is loaded with cold antiprotons. The trap has a diameter of 12 mm, the antiproton plasma a diameter of ≈ 1 mm FWHM. Then a bunch of $8 \cdot 10^7$ positrons is shot onto the production target, positrons are converted into positronium and re-emitted in the general direction of the production trap. This results in an extremely intense γ ray burst that, despite the limited sensitivity of FACT to γ rays completely saturates the detector (see as a reference the plot in the opening image of appendix D), followed by a second smaller burst due to positronium annihilation that cannot be detected through FACT (being it completely overcome by the previous more intense burst) but can be seen by the scintillating detectors (it is the bump in the red curve of figure 3.13). If lasers are shot with the right timing with respect to the positron shot then antihydrogen can be produced in the trap; antihydrogen is not bound by the production trap's electromagnetic field so it will reach the trap wall and annihilate there.

The temperature of the produced $\bar{H}$ is dominated by the trapped antiproton temperature and is therefore expected to be around 150 K. Being $\bar{H}$ neutral we expect it in first approximation[3] to fly in a straight path in the production trap, oblivious of the electromagnetic fields. Figure 4.1 shows the expected annihilation rate of a cloud of antihydrogen produced near the axis of the production trap with a temperature of 150 K which annihilates on the trap walls after flying in a straight path; by integrating the area under the profile we can see that over 90% of the produced $\bar{H}$ reaches the trap walls in the $0 \div 12$ µs range.

We determined that the average lifetime of $\bar{p}$ in the production trap exceeds 1000 s; this means that in any given time window of 12 µs we expect less than 0.005 annihilations to take place on average. Upon this background an excess signal of a few $\bar{H}$ annihilations should clearly emerge; still to be able to determine this excess to be due to $\bar{H}$ we need to exclude the possibility of the mean life of antiprotons in the production trap being locally decreased by the production and excitation of positronium. The implantation of positrons into the NCP target requires switching of electrodes and manipulation of a charged plasma near the production trap, which can perturb it; the excitation of positronium requires shooting lasers with a peak power over 100 MW which can evaporate materials in the production region, locally degrading the vacuum. The two phenomena can be studied separately, by shooting either solely the laser pulse or the positron bunch and determining whether the presence of one of the two (which is not enough to cause the production of a significant

[3] Of course a precise computation needs to take into account the effects of Stark acceleration.

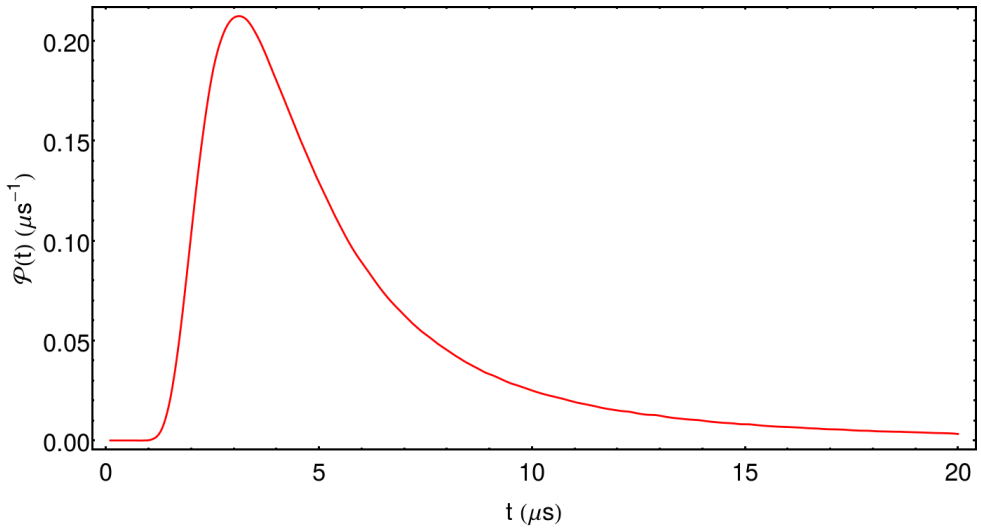


Figure 4.1: Probability for an $\bar{\mathrm{H}}$ produced at time 0 to reach the production trap walls at time t. The $\bar{\mathrm{p}}$ cloud has been assumed to have a temperature of 150 K and a gaussian radial profile with $\sigma =$ 0.5 mm. Straight trajectories have been simulated (no Stark effect acceleration has been considered).

amount of antihydrogen) increases locally the $\bar{\mathrm{p}}$ background. Main assumption of this technique is that no background increases due to synergy between the operation of the laser and the NCP converter. We do not expect there to be any, but depending on the degree of certainty required to determine whether $\bar{\mathrm{H}}$ production took place or not, independent proof of $\bar{\mathrm{H}}$ production might be required.

Another element that can be employed to effectively discriminate between $\bar{\mathrm{H}}$ and $\bar{\mathrm{p}}$ annihilations is the spatial collocation of the annihilation vertexes inside of the production trap. The production trap has a radius of 6 mm which greatly exceeds the $\bar{\mathrm{p}}$ plasma radius. We expect most of the $\bar{\mathrm{p}}$ annihilations to take place close to the trap axis, whereas $\bar{\mathrm{H}}$ atoms will likely escape the electromagnetic trapping and annihilate onto the trap walls. Let's fix a cylindrical coordinate system, having the axis in common with the production trap and conventional coordinate names r for the radial, ϕ for the azimuthal and z for the axial. By determining the radial coordinate r of the annihilation vertexes we can discriminate $\bar{\mathrm{H}}$ and $\bar{\mathrm{p}}$ annihilations. A chance to do this might arise from employing the FACT detector.

4.2 The FACT detector

Surrounding the production trap, AE$\bar{\text{g}}$IS features the Fast Annihilation Cryogenic Tracker (FACT) detector[124]. The FACT detector consists in 796 scintillating fibers[4] with a diameter of 0.5 mm wound in circles that share the production trap axis. Fibers are wound up to four different radii and fibers with the same winding radius are mounted back-to-back with a 0.2 mm gap separating them, forming an almost complete coverage of a cylindrical surface surrounding the production trap. The four layers of fibers are divided into two macrolayers with radii 70 mm and 98 mm; a macrolayer is made up of two layers wound at radii that differ by:

$$\Delta R \;=\; \sqrt{4r^2 \;-\; \left(r + \frac{\delta r}{2}\right)^2} \;=\; 0.8\,\text{mm} \tag{4.1}$$

and staggered along the z axis (figure 4.2 shows in section the placement of the two layers of fibers in a macrolayer). In the previous formula $r = 0.5\,\text{mm}$ is the radius of a fiber and $\delta r = 0.2\,\text{mm}$ is the gap left between subsequent fibers within a layer. Scintillating fibers are coupled with clear fibers which convey light to silicon MPPCs[5]; signals from the MPPCs are amplified and routed into fast discriminators to produce a digital TTL signal that is then recorded through an array of 17 FPGA development boards[6]. The FPGA boards are read out through ethernet connections by a control computer. A great deal of work has been done to prepare the FACT detector for data taking during the 2016 and 2017 antihydrogen runs, and a good portion of it consisted in developing software capable of efficiently the calibrate and read out of the detector; a detailed description of how the technical challenges posed by this task have been overcome can be found in appendix D.

Data from the FACT detector consists in the timings for each rising edge an subsequent pulse length, both recorded with a resolution of 5 ns. Whatever physics information we want to extract from the FACT detector needs to be extracted from this.

4.3 Vertexing in FACT

As previously stated we are interested in the reconstruction of the r coordinate of annihilation vertexes in the production trap. This employment exceeds the original FACT design which was to detect the z position of annihilation vertexes lying close to the trap axis. This poses severe limitations to our

[4] Kuraray SCSF-78 M
[5] Hamamatsu S10362-11-100C
[6] Xilinx Spartan-6

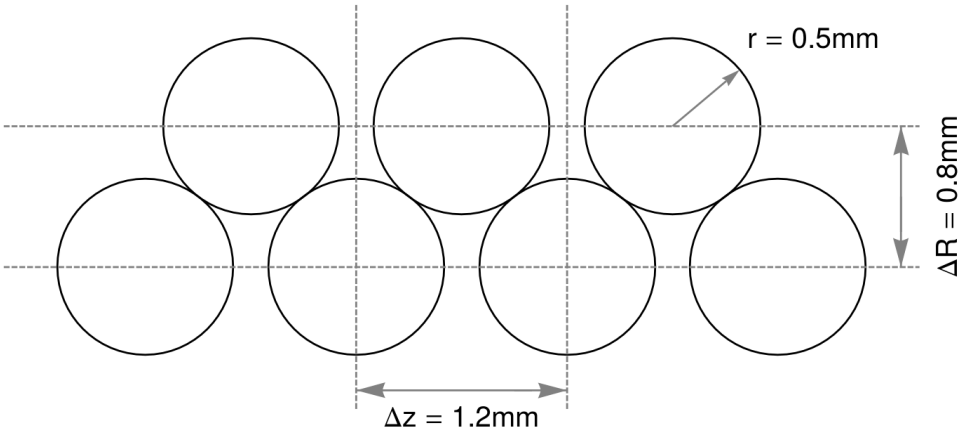


Figure 4.2: Sectional view of fibers from two adjacent layers displaying the staggered arrangement of the fibers of FACT.

work, the first of which is the mathematical impossibility to reconstruct the azimuthal angle of annihilation vertexes due to the detector being completely degenerate along the ϕ coordinate.

Tracks in FACT can be conveniently approximated as straight lines. Charged pions from antiproton decays in the 1 T magnetic field are indeed expected to undergo a helicoidal motion, still in the worst case scenario we can expect them to have a trajectory with a curvature over 1 meter. This means that, at the scale of the FACT detector, trajectories will deviate by less than one millimeter from a straight path; this deviation is at the edge of the detector's resolution. The reconstruction techniques presented hereafter can be perfected by adding corrections that account for the curvature of the trajectories; nonetheless these corrections contribute nothing to neither the exposition nor to the conclusions.

To reconstruct the r coordinate the conventional approach would be to collect closely-timed rising edges into coincidences, then to gather spatially adjacent fibers firing inside of the same coincidence into clusters and then to fit tracks onto these clusters. Here the limitation given by the degeneration over the ϕ coordinate strikes us the hardest: when projected into the $r - z$ plane linear tracks turn into hyperbolae in the form:

$$r^2 = (a + bz)^2 + (c + dz)^2 \tag{4.2}$$

which, as opposed to linear forms require 3 points to be determined. Since we have only two macrolayers of fibers we can at most hope for two clusters to fit the straight line over. If we impose the vertex of the hyperbola to sit

onto the trap axis we make the problem be determined with just two clusters, otherwise it would seem impossible to determine the r coordinate of vertexes in FACT.

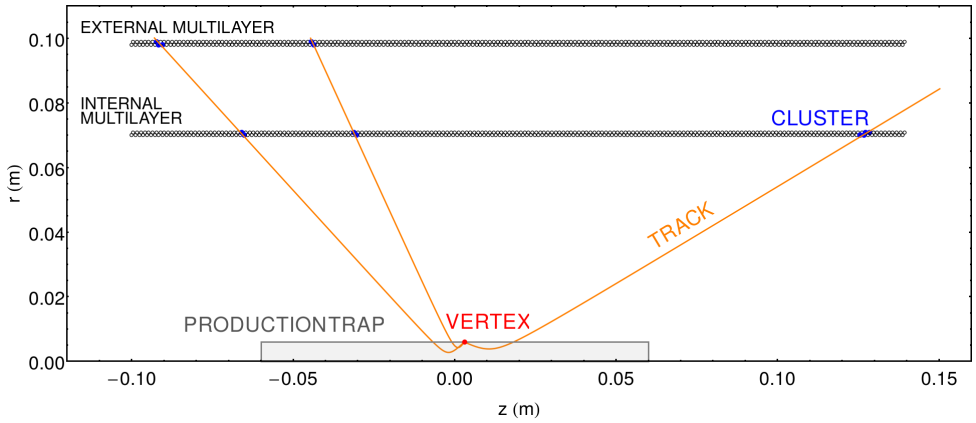


Figure 4.3: From a single annihilation vertex on the wall of the production trap three γ rays emerge. The straight trajectories appear, due to being projected onto the $r - z$ plane, as hyperbolae.

Actually whenever we are attempting the conventional vertex reconstruction technique described above we are ignoring a lot of the information that the data is conveying. First of all the shape of the clusters, although it does not determine by itself the position of the vertex, is influenced by it (see, for example, figure 4.4). Secondly even though each pair of clusters cannot determine the position of a vertex inside of the trap, the probability of clusters being detected at certain positions is not independent on the vertex radial position. Finally although the proposed method attempts the reconstruction of vertexes using single coincidences, data accumulated from large amounts of coincidences, even across different runs, yields additional information on the radial distribution of the annihilation vertexes.

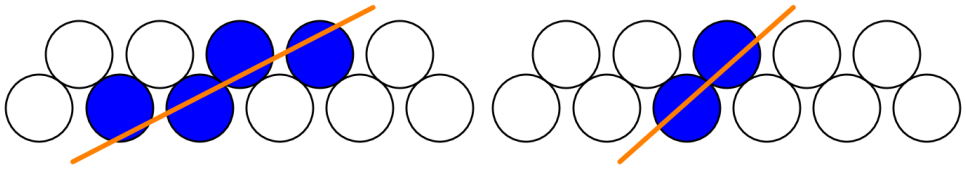


Figure 4.4: The shape of clusters of activated fibers encodes some information about the shape of the particle's trajectory

We'll describe hereafter an ongoing attempt to educe all the information held by the FACT data about the radial distribution of annihilation vertexes in the production trap during the operation of the AE$\bar{g}$IS apparatus.

4.4 Bayesian vertexing

The last decade of the 20th century has seen the re-emergence of the Bayesian approach to probability and flourishing of its application in the field of physics experiments data analysis[125]. A solid treatment of the subject is given by D'Agostini[125]; a more brief introduction to the practicalities of Bayesian inference is given by the same author in [126, 119].

In brief, whenever we run an experiment we measure observables whose expected outcome is, in our model of the experiment, correlated with a hidden variable that cannot be detected directly. What we are interested in is to determine the probability that a certain observation was caused by the hidden variable[7] having a certain value rather than another. In the specific case of FACT we observe the synchronous stimulation of a set of fibers and we would like to predicate the probability of the event being caused by an annihilation whose vertex has a certain location.

Usually the formulation of the model of the experiment allows the easy computation of the probability $\mathcal{P}(O|H_n, K)$ of a certain observable O to manifest, given that the hidden variable has a certain value H_n and all of the knowledge K we hold about the detector and experiment physics. What we are interested in determining, on the contrary, is the *Posterior* probability $\mathcal{P}(H_n|O, K)$ of the hidden variable being H_n given that we observed O. The Bayes theorem allows us to do so as:

$$\mathcal{P}(H_n|O, K) \;=\; \frac{\mathcal{P}(O|H_n, K)}{\mathcal{P}(O, K)} \cdot \mathcal{P}(H_n, K) \tag{4.3}$$

two new factors appeared in equation 4.3. $\mathcal{P}(O, K)$, appearing at the denominator, holds little significance to us since, as we will see, can be eliminated through normalization. $\mathcal{P}(H_n, K)$ is called the *Prior* distribution and describes the knowledge that we hold about the hidden variable (such as the position of vertexes in FACT) before performing the measurement. The final result will depend from the value of $\mathcal{P}(H_n, K)$, and rightly so. It is reasonable to say that whenever this dependency maps a reasonable span of *Prior* probabilities into similar *Posteriors* then that the measurement results are well defined, on the other side if reasonable *Priors* map into widely different *Posteriors* then the measurement is prone to interpretation and further data should be collected to narrow the result.

It is wise to better interpret the meaning of equation 4.3 to consider the same relation applied to another possible value H_m for the hidden variable and then to compute the ratio of the two equations. If we do so we obtain:

[7] In our case the position of the annihilation vertex.

$$\frac{\mathcal{P}(H_n|O,K)}{\mathcal{P}(H_m|O,K)} = \frac{\mathcal{P}(O|H_n,K)}{\mathcal{P}(O|H_m,K)} \cdot \frac{\mathcal{P}(H_n,K)}{\mathcal{P}(H_m,K)} \tag{4.4}$$

which is of easier interpretation. On the left we have the *Posterior* assessment of how likely we expect the hidden variable to have value H_n as opposed to H_m, on the right we have the *Prior* assessment of the same quantity and in the middle we have the *Likelihood* ratio that represents the result of our measurement. We can use the relation in equation 4.4 to compute the value of the posterior distribution over all of the possible values of H by first computing the product

$$\mathcal{P}(O|H_n,K) \cdot \mathcal{P}(H_n,K) \tag{4.5}$$

for every possible H_n. We know from 4.4 that the posterior probability is proportional to 4.5 so we can write the posterior probability by normalizing 4.5:

$$\mathcal{P}(H_n|O,K) = \frac{\mathcal{P}(O|H_n,K) \cdot \mathcal{P}(H_n,K)}{\sum_i \mathcal{P}(O|H_i,K) \cdot \mathcal{P}(H_i,K)} \tag{4.6}$$

Now, after each measurement (in our case after analyzing a single coincidence) the obtained *Posterior* will become the new best assessment of $\mathcal{P}(H_n,K)$ being O now absorbed into K. We can therefore pick a *Prior* distribution, apply the Bayesian inference to transform it into a *Posterior* that we will then feed as the *Prior* of the next iteration or, equivalently, we can multiply together the likelihoods of all of the observed events to obtain a single *Likelihood* distribution that is capable of propagating the original *Prior* (the one we held before even starting the measurements) into the final *Posterior* (the one that we hold after taking into account all of the data gathered from every recorded coincidence). We will follow the latter strategy.

4.5 A toy model

To better understand the methodology employed to perform the reconstruction of vertexes we present here a toy model that makes use of an imaginary detector. Let's imagine a square box in a two dimensional plane, in the box is placed a radioactive source which occasionally emits a particle in a random direction, the emitted particle travels in a straight line until it reaches a one of the walls of the box. The walls of the box are divided each into 10 segments of equal length. Each segment is a particle detector capable of determining whether it has been hit by a particle or not; all of the segments, considered together,

will therefore constitute a 40-pixel detector. Let's see if we can determine the position of the radioactive source using Bayes theorem to update our best guess on its position after each particle detection.

To do so numerically we start by tassellating the space inside of the detector in a 10×10 grid of cells: we won't reconstruct vertexes positions below this resolution. Then for each cell n we compute the probability $\mathcal{L}(n, m)$ that a particle coming from a source randomly placed inside of that cell will hit the pixel m; this can be done in general by means of Monte Carlo integration or, in this specific case, by systematic sampling which is convenient due to the integration domain being only three dimensional. Since we will need it afterwards let us write down this integral now before adding any further complexity to the system:

$$\begin{aligned} \mathcal{L}(n,m) &= \mathcal{P}(O_m|H_n) \\ &= \frac{1}{\int_\Phi 1\, d\phi} \cdot \int_\Phi \chi(H_n, O_m, \phi)\, d\phi \\ &= \frac{1}{2\pi \Delta x \Delta y} \cdot \int_{\alpha=0}^{2\pi} \int_{x=x_n}^{x_n+\Delta x} \int_{y=y_n}^{y_n+\Delta y} \chi_P(O_m, x, y, \alpha)\, d\alpha\, dx\, dy \end{aligned} \tag{4.7}$$

Let us start by commenting the second integral. $\chi_P(O_m, x, y, \alpha)$ is a function that evaluates to one if a particle emitted in the point (x, y) with angle α hits the m-th pixel, to zero otherwise; x_n and y_n are the coordinates of one corner of the n-th grid cell, $\Delta x = \Delta y$ the cell sizes. The form written above is the generalization of the procedure we are adopting in this specific case: the outcome of the measurement depends on several variables that we are not interested in reconstructing ($x - x_n$, $y - y_n$ and α in this case), the value of these variable is identified by a point ϕ in a space Φ. We can then write the conditional probability of observing O_m given the hidden variable H_n as the integral of a function $\chi(H_n, O_m, \phi)$ that evaluates to 1 if the value ϕ for the other variables causes O_m to be observed when the hidden variable is H_n and 0 otherwise.

If we take n to be the row index and m to be the column index of a matrix, then $\mathcal{L}(n, m)$ is a right stochastic matrix, i.e.: it contains nonnegative values and the content of each of its rows sums to 1. To apply Bayes theorem we need to do nothing more than to read the matrix by its columns. Each column of the $\mathcal{L}$ matrix is a non normalized distribution on the possible values for the hidden variable H, specifically the *Likelihood* of the observable O_m associated with the m-th column.

We then proceed to test the reconstruction algorithm. First of all we choose randomly a point P inside of the detector. Then we generate randomly-directed particles originating from point P and compute which pixel gets activates

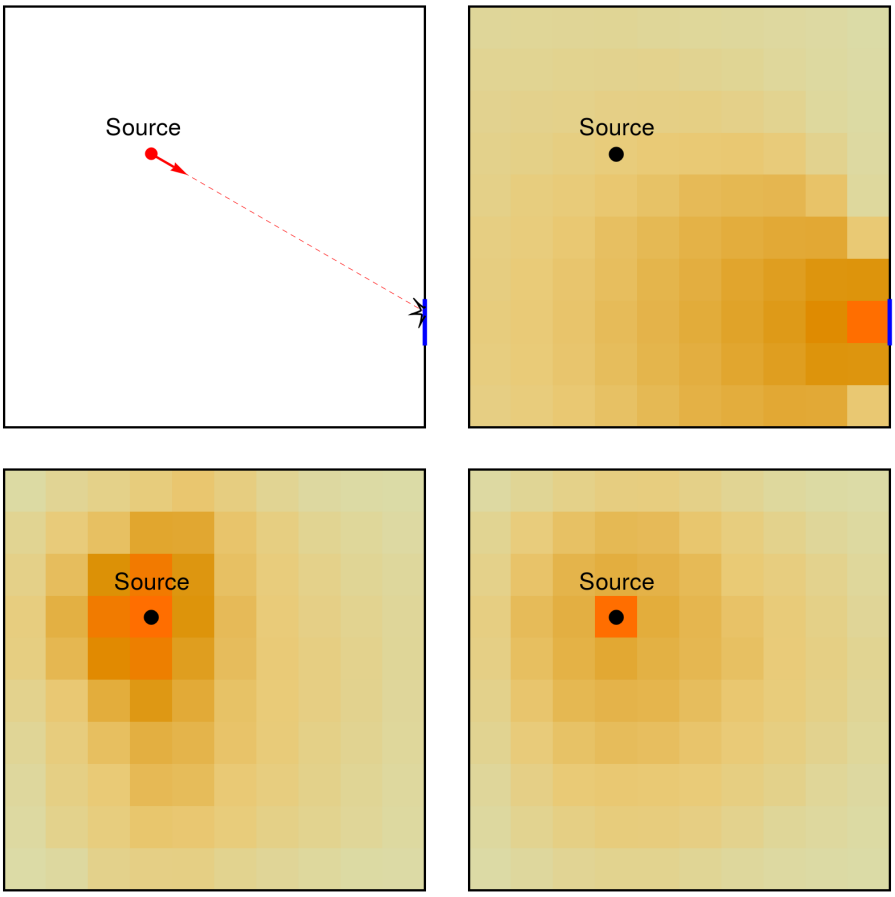


Figure 4.5: A toy model to attempt vertex reconstruction using Bayes theorem. On the top left the experimental setup: a source placed inside of the box emits a particle that is detected by a pixel placed on the wall (in blue). On the top right the likelihood distribution associated with the activation of the specified pixel. On the bottom left the likelihood distribution given by 35 combined observations. On the bottom right the combined likelihood distribution given by 300 combined observations.

by each of these particles. Then for each pixel activation we extracting its associated *Likelihood*. Finally by multiplying together the so obtained *Likelihoods* we compute the likelihood relative to the detection of all of the simulated particles. If the reconstruction is successful this will result in a distribution that features a peak that, as the number of detected particles increases, converges onto the source location. Figure 4.5 shows the toy model setup, the shape of a single event *Likelihood*, of a 35-event *Likelihood* and of a 300-event *Likelihood*.

4.5.1 Missing hypotheses problem

There are two characteristics that are best observed in the toy model that should be addressed before proceeding further. The first is that this procedure is very effective in distinguishing between different hypotheses for the hidden variable and returns results that are very solid not only in their final value but also in the estimation of their uncertainty. What must be cleared is that, given a reasonable *Prior* the procedure computes the relative probability of all of the proposed hypotheses for the hidden variable; were the value for the hidden variable not sufficiently close to any of the proposed ones the method will likely converge to the best hypothesis among the available ones.

Let's consider what happens in the toy model if the source point is not fixed in one point but randomly selected among two possible locations. If the two source points have equal probability of producing a particle then the method will converge, much more slowly, to a point located approximately between the two locations. If, on the other side, one of the two points is selected with a much higher probability than the other (such as 10 to one), then the inference will converge onto its location and ignore the other source.

At the same time we can ask the method to correctly distinguish between sources of arbitrary shape as long as we list them as possible values of the hidden variable. In figure 4.6 we can see how the method is able to distinguish between two extended sources as long as they are correctly presented as possible values for the hidden variable.

In general it is not possible to employ all of the possible combinations of hidden variable values (in the first formulation of the toy model a 10×10 grid results in $2^{100} - 1 \approx 1.27 \cdot 10^{30}$ possible combinations) so the physics of the experiment must be employed to narrow down the hypotheses pool before performing the regression.

In the case of FACT, a possibility might be to ask the reconstruction to distinguish between annihilations happening on the trap axis with a distribution similar to that of the antiproton plasma and annihilations happening on the trap wall with a distribution computed via Monte Carlo simulation predicting where we expect the $\bar{\mathrm{H}}$ to annihilate. Of course combinations of the two, along with realistic competing hypotheses should be added to the

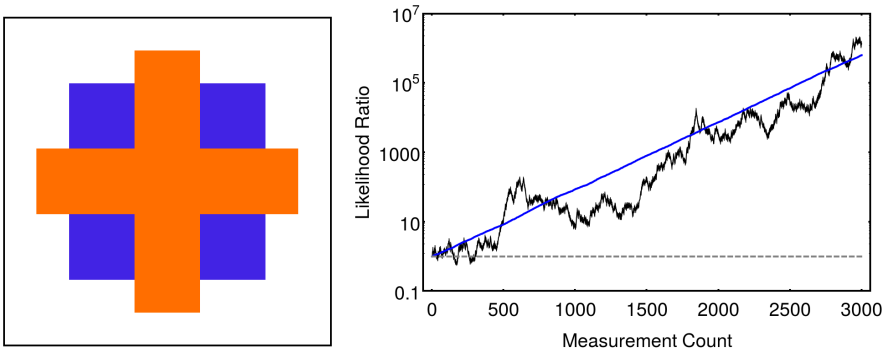


Figure 4.6: We employ one of the two extended sources displayed in the left panel to generate events and employ the Bayesian regression to determine which one of the two is being used to generate events. On the right we see the evolution of the likelihood ratio (black) as a function of the number of measurements performed. In blue the average evolution of the likelihood ratio computed averaging (in log scale) 1000 instances of the experiment.

pool of possible values for the hidden variable.

4.5.2 Multiple correlated emissions

The proposed model holds even in the case of a source that emits several particles whose emission angles are correlated. Let's consider the slightly different toy model in which the source placed in the detector described above emits three particles at a time whose trajectories form angles of 120°.

This variant is interesting for a couple of reasons. The first is that since now every decay can excite up to three pixels we have a pool of 11480 potential observables; of these only 3959 manifest due to the constraints posed by the process physics. This is interesting on one side since the fact that only a portion of the potential observables can actually manifest will be crucial in the real world scenario to tame the exponential combinatorial explosion of possible sets of activated fibers; on the other side it raises the problem of "impossible" outcomes, that is observables that are not admitted by our model. Columns of the $\mathcal{L}(n, m)$ matrix relative to observables that are non admitted by our model contain only zeroes and were the *Prior* to be multiplied by them it would yield a uniformly zero non normalizable *Posterior*. Rightly so since update of probabilities cannot be performed when impossible events are observed. In the non ideal world such events, unfortunately, can occur due to an incomplete model of the experiment or of the apparatus noise; our real implementation, as we will see, will need to address how to manage "impossible" outcomes.

Another reason this variant is interesting is that information encoded in

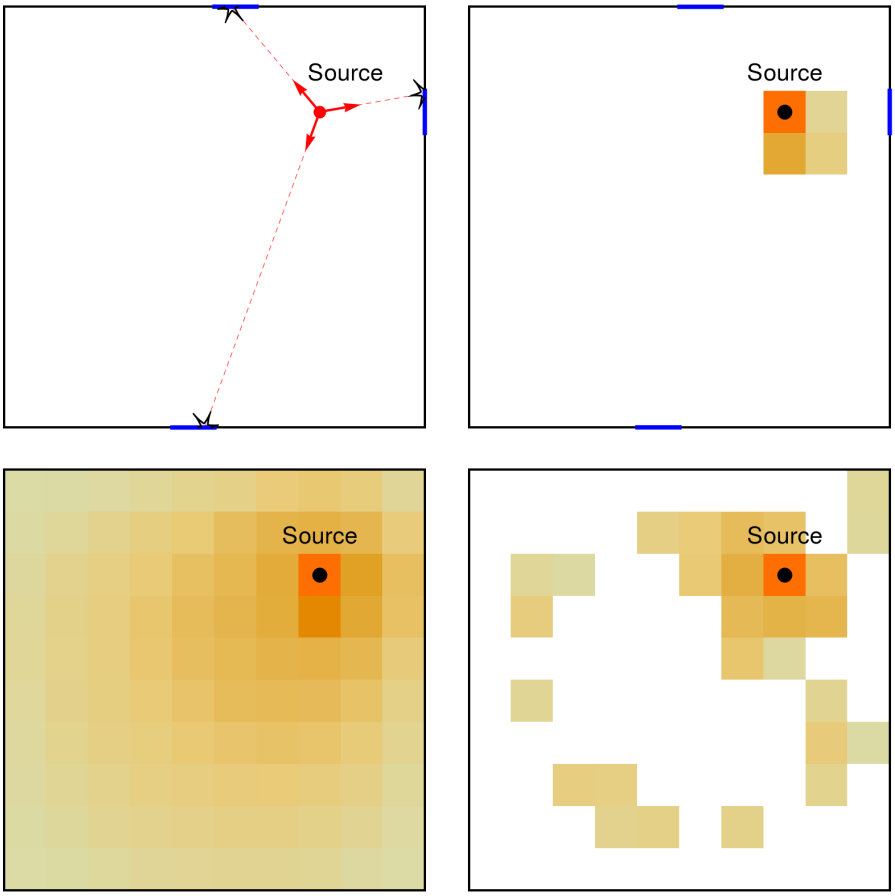


Figure 4.7: We employ now decays that emit three 120°-angled particles (see top-left panel). On the top right we see the likelihood of a single event in this scheme, which allows for direct vertex reconstruction. On the bottom left we see the likelihood of 30 events of this model reconstructed while ignoring completely the correlations among emission angles. On the bottom right the likelihood associated with 10 events reconstructing using a model that accounts for correlations between two of the three emitted particles. White cells denote H_ns for which the associated *Likelihood* has been determined to be exactly zero; as anyone accustomed to work with Bayes regression, models yielding zeroes can become problematic and need some careful handling when employed.

the observable associated with a single decay is sufficient to locate a vertex inside of the detector. This characteristic manifests in the *Likelihoods* taking a shape similar to a delta distribution. This is an instance in which traditional vertex reconstruction techniques could be well employed. Instead we will take a chance to show two characteristics of the vertex reconstruction based on Bayes regression.

The first is quite intuitive: we can employ a model that does not describe correlations between emission angles to build the $\mathcal{L}(n,m)$ matrix and, although less efficiently, reconstruct vertexes inside of the detector. The results are shown in figure 4.7. More generally we can note that our observable is influenced by the value of the hidden variable H_n and by the value of a set of other variables ϕ in the space Φ, with the notation introduced in equation 4.7. The shape of the Φ region can impose correlations between these variables (such as the emission angles always differing by a fixed amount). If we loosen our model by removing these correlations we, in general, reduce the convergence speed of the method but do not bias the final result.

The next characteristic we can exhibit is that we can write an approximation of the likelihoods $\mathcal{L}_3(n,\{m,p,q\})$ relative to three-particle decays with uncorrelated emission angles using the *Likelihood* matrix $\mathcal{L}_1(n,m)$ for single particle decays. The entries of the three-particle decays *Likelihood* matrix are, similarly to equation 4.7:

$$\begin{aligned}
\mathcal{L}_3(n,\{m,p,q\}) \;&=\; \mathcal{P}(O_{\{m,p,q\}}|H_n) \\
&\approx \frac{1}{\left(\int_\Phi 1 d\phi\right)^3} \cdot \iiint_\Phi \chi(H_n,O_m,\phi_1)\,\chi(H_n,O_p,\phi_2)\,\chi(H_n,O_q,\phi_3)\,d\phi_1\,d\phi_2\,d\phi_3 \\
&= \frac{1}{\left(\int_\Phi 1 d\phi\right)^3} \cdot \left[\int_\Phi \chi(H_n,O_m,\phi_1)d\phi_1\right] \cdot \left[\int_\Phi \chi(H_n,O_n,\phi_2)d\phi_2\right] \cdot \left[\int_\Phi \chi(H_n,O_q,\phi_3)d\phi_3\right] \\
&=\; \mathcal{L}_1(n,m)\,\cdot\,\mathcal{L}_1(n,p)\,\cdot\,\mathcal{L}_1(n,q) \equiv \mathcal{L}_{Approx}
\end{aligned} \tag{4.8}$$

This approximation is already extremely good, with the relative difference between the approximated and exact *Likelihood* $(\mathcal{L}-\mathcal{L}_{Approx})/\mathcal{L}$ ranging typically under 2% with ocasional spikes that can reach up to 30% (see figure 4.8) but that do not impair the reconstruction. The distance between the exact and the approximated *Likelihood* can be improved by reducing the grid spacing.

Unluckily removing the correlations between the emission angles of the three particles greatly reduces the efficiency of the reconstruction. Let's assume that we have at our disposal the *Likelihoods* for the one-particle decay $\mathcal{L}_1(n,m)$ and for correlated two-particle decays $\mathcal{L}_2(n,\{m,p\})$. We can use them to compute directly an approximation of the *Likelihoods* of a model in which three particles are emitted: two of them have correlated emission angles

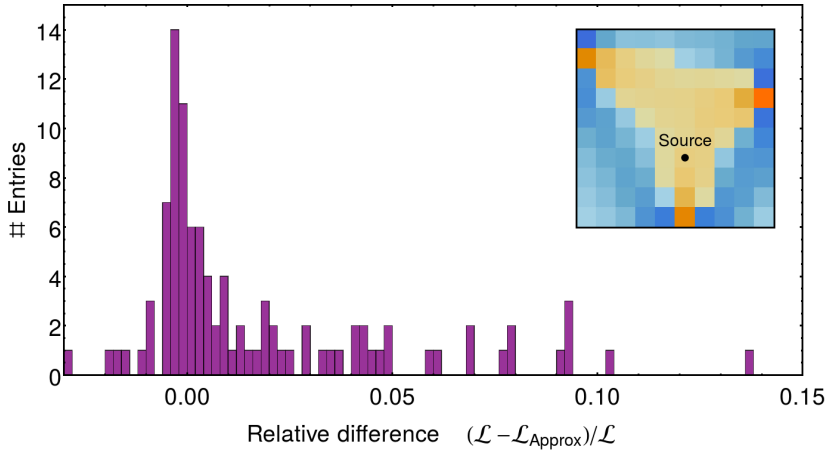


Figure 4.8: An example *Likelihood* for the uncorrelated emission of three particles has been computed exactly and approximated by multiplying *Likelihoods* computed for single particle emission. The plotted curve is the relative difference $(\mathcal{L}-\mathcal{L}_{Approx})/\mathcal{L}$ over all of the hidden variable values H_n. In the inset the difference between the likelihoods normalized separately to have a total integral of 1.

while the third is emitted at a random angle. Let Φ_1 be the single particle other variable space (that in this case comprises the vertex position inside of the cell and an emission angle), and Φ_2 be the correlated two-particle hidden variable space (that in this case comprises the vertex position inside of the cell and two emission angles that differ by exactly 120°). Let χ_1 be the indicatrix function for one particle, and χ_2 the indicatrix function for two particles.

$$\begin{aligned}
\mathcal{L}_3^*(n,\{m,p,q\}) \;&\approx \\
&\approx \frac{1}{3\cdot\left(\int_\Phi 1\,d\phi\right)^3}\cdot\int_{\Phi_1}\int_{\Phi_2}\chi_1(H_n,O_m,\phi_1)\cdot\chi_2(H_n,O_{\{p,q\}},\phi_2)\;+\\
&\qquad\qquad\chi_1(H_n,O_p,\phi_1)\cdot\chi_2(H_n,O_{\{m,q\}},\phi_2)\;+\\
&\qquad\qquad\chi_1(H_n,O_q,\phi_1)\cdot\chi_2(H_n,O_{\{m,p\}},\phi_2)\;d\phi_1\,d\phi_2\\
&= \frac{1}{3}\Big[\mathcal{L}_1(n,m)\mathcal{L}_2(n,\{p,q\})\;+\;\mathcal{L}_1(n,p)\mathcal{L}_2(n,\{m,q\})\;+\;\mathcal{L}_1(n,q)\mathcal{L}_2(n,\{m,p\})\Big]
\end{aligned} \tag{4.9}$$

That is we average the *Likelihoods* relative to every possible attribution of the three excited pixels to one or the other of the two underlying emission processes. The effect of restoring partially the correlations among emitted particles in the model is to greatly increase the effectiveness of the reconstruction

process. In figure 4.9 the average *Likelihood* ratio between the cell containing the source and the rest of the grid is computed using the correlated three particle model, the fully uncorrelated model and this last one. The result is that in the fully correlated model the *Likelihood* ratio typically diverges at the second measurement (that is: only the cell containing the source is compatible with both observations). The fully uncorrelated model causes a slow exponential divergence that obtains a *Likelihood* ratio of 100 within a hundred observations. The hybrid model displays an abysmally faster exponential explosion which typically then around the 35th measurement diverges to infinity as the fully correlated model did.

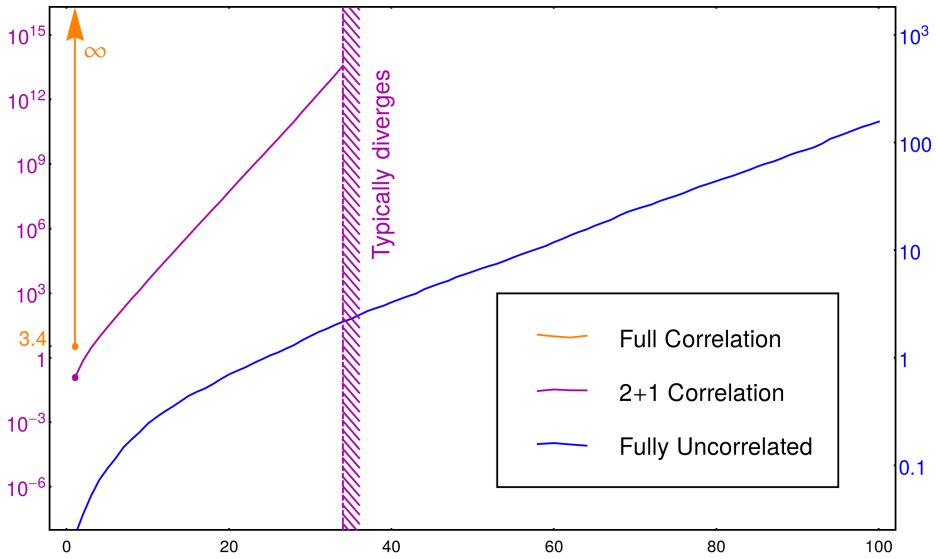


Figure 4.9: The typical *Likelihood* ratio for the fully correlated, fully uncorrelated and hybrid models are compared. The mean of several runs is displayed. Since in this model the value of the *Likelihood* can be equal to zero in some cells, the *Likelihood* will eventually diverge after enough measurements have been taken into account. In our case the fully correlated model typically diverges after the second measurement, while, where the models typically yield ∞ the plot have been interrupted

4.6 The real world scenario

Toy models are usually easy and fun to run since we design them to be so; the same cannot be said for the real world scenario. Applying the exact bayesian regression to reconstruct vertexes in FACT is computationally unfeasible, nonetheless we can aim to approximate the technique closely enough to

perform a satisfactory vertex reconstruction.

The main obstacle towards performing the complete reconstruction is the potentially enormous observable space. The 796 fibers of the detector can be potentially activated in $2^{796}-1 \approx 4.17 \cdot 10^{239}$ different ways; luckily most of these potential observables never manifest in our models. Moreover a great deal of the potential observables appear so rarely that their combined probability to manifest is small enough that they can be ignored altogether while maintaining the capability to reconstruct most of the events that actually appear in a real case scenario.

A Monte Carlo simulation program based on Geant has been developed to accurately generate realistic FACT events. Unfortunately it is too slow to be effectively employed in our tests so its usage needs to be postponed to a latter stage of development. Instead we developed a simpler and much faster Monte Carlo tool that takes into account only the geometric shape of the fibers of FACT and is restricted to determining which fibers are intersected by a straight half line coming from a vertex located inside of the production trap. This model can be run on a laptop at a rate of about 1 Mevent/s which grants us some elbow room to run tests.

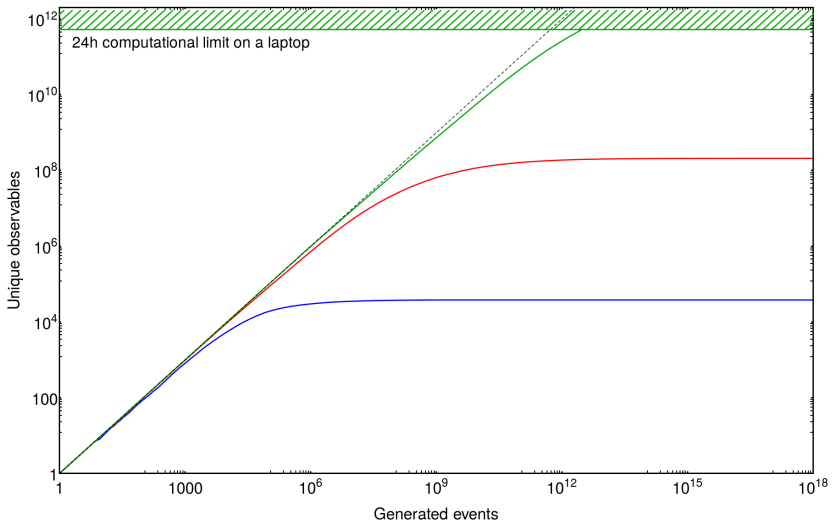


Figure 4.10: The amount of unique different observables produced by the purely geometric Monte Carlo generator for FACT for one-particle decays (blue), two-particles decays (red) and three-particle decays (green). In dashed black the identity function is shown as a reference.

We partitioned the production trap into a 6×82 grid with 1 mm spacing and generated events consisting of up to n particles originating from the same point (randomly placed inside of a cell of the aforementioned grid) and

traveling in uncorrelated random directions. We try to build with these events a *Likelihood* matrix, which will be possible only if such matrix is sparse enough.

In figure 4.10 the number of non null columns in the *Likelihood* matrix is shown as a function of the generated events. We can see that for one-particle events the entire matrix can be easily built, for two-particle events generating the matrix requires some effort but still lies within the realm of possibility whereas three-particle or higher events would need an impractical amount of computational power to be mapped into a likelihood matrix.

The model described up to this point is unlikely to ever match more than a small subset of the recorded FACT events. Two fundamental elements are missing: noise and fiber efficiency.

Of them the less problematic might be noise. We calibrate the FACT detector digitizer thresholds so that each fiber has a dark count of about 100 Bq. This means that in any 5 ns window we expect approximately $4 \cdot 10^{-4}$ dark counts over the entire detector; we usually define coincidences as rising edges detected within 20 ns or less so we expect less than 0.1% of coincidences[8] to contain dark counts. This might not hold true since we recently observed to our demise correlation between dark noise and signal rate.

The main obstacle plaguing the employment of the above described model is the lack of account for the efficiency of the fibers. As of today we do not have an accurate estimate of this parameter (which of course would be also dependent on the energy of the incoming particle) but early estimates place it closer to 0.5 than to 1. This means that if we were to include this behavior inside of the model employed to build the *Likelihood* matrix we would severely cripple any chance we might hold of employing two-particle events to perform the reconstruction. This is because accounting for inefficient fibers inside of the model substantially increases the amount of allowed observables, since the model will have to include all of the possible configurations of the missing detections by the fibers.

Instead we can account for noise and fiber efficiency during reconstruction. As we will see this approach will have the advantage of relieving us of fixing the noise and efficiency parameters during the (costly) computation of the likelihood matrix and will allow us, instead, to produce parametric results that are polynomially dependent on these two parameters as we will detail in the rest of this section.

[8]The procedure described here is not complete, there exists a little contribution to the actual dark count in coincidences deriving from the selection bias: i.e.: the presence of dark counts makes a coincidence more likely to pass any filter that we employ to discriminate coincidences from unwanted events. Due to the low rates even accounting for that does not change substantially the result.

4.6.1 Combinatoric *Likelihoods*

Let's consider the case in which we need to account for the recording of a coincidence where N fibers experienced a rising edge. Let's call these fibers f_i and the set $\{f_i \,|\, \forall\ 1 \leq i \leq N\}$ as F. We need to produce, based on the one-particle and the two-particle *Likelihood* matrices $\mathcal{L}(n,m)$ (notation from section 4.5), an approximated likelihood $\mathcal{L}^F(n)$ for this event. The approximate *Likelihood* will have to account for the event being possibly generated by more than two particles, for the limited efficiency of the fibers and for the possible presence of dark counts. We will show in the following how this can be done if only the one-particle *Likelihood* matrix is available; this case can be extended to make use also of the two-particle likelihood but we don't consider the technicalities of this extension to be neither interesting nor innovative.

The model upon which is based the *Likelihood* we are trying to approximate is not completely defined unless we provide an overall probability for the emission of q particles whenever a decay event takes place. Let $\mathcal{N}(q)$ be the probability that a decay taking place inside of the detector will result in the emission of exactly q particles. Let's also assume $\mathcal{N}(q)$ is normalized so that:

$$\sum_{q} \mathcal{N}(q) \ = \ 1 \tag{4.10}$$

In practice we can assume that $\mathcal{N}(q)$ will be exponentially depressed at high q (after we surpass the kinematic limitations of the physics of the decay of $\bar{\mathrm{H}}$, the only way more particles can be produces is if two decays take place at the same point at the same time which gets exponentially less probable). In practice we can assume that after the very first few values, $\mathcal{N}(q)$ can be safely assumed to be zero, let's call this limit Q and therefore $\mathcal{N}(q) = 0\ \forall\ q > Q$

We will assume that a single decay emitting up to Q particles has excited the F set of fibers. We start by considering all possible subdivisions of F into up to Q (not necessarily disjoint) subsets such that the union of all of the subsets in a single subdivision is again F. Let S be this partitioning set, we can write it as:

$$S \ = \ \left\{ \ S_j \ \in \ \bigcup_{q=1}^{Q} [\,\mathcal{P}(F)\,]^q \ \ \middle| \ \ \bigcup_{k=1}^{K_j} S_{j,k} \ = \ S_j \ \right\} \tag{4.11}$$

Were S_j are the elements of the partition set, let $J = |S|$ be the total size of S. $S_{j,k}$ are the elements of the j-th partition, K_j the number of sets in the j-th partition. In our construction we are going to postulate that each set of fibers $S_{j,k}$ has been excited by a single particle.

Without entering into detailed computations, we can give come coarse bounds on J as $\frac{1}{Q!}(2^Q-1)^N \leq J \leq (2^Q-1)^N$ which for realistic numbers can be excessively high (for $Q=6$ and $N=8$, $3.44\cdot 10^{11} \leq J \leq 2.49\cdot 10^{14}$). This is the first of several instances in which we encounter combinatoric enumerations that cannot be realistically computed by sheer brute force. Indeed when we will cycle over S in equation 4.12 we expect the vast majority of the elements of S to give no contribution. Computation over the S set is manageable as long as we employ some smart technique. to curb the problem; an example of which is building S recursively like it was a tree exploration many branches of which we trim down early as soon as we are able to determine that any of its sub-branches will lead only to null contributions and many others we compute only once since we realize that from a certain point onward their structure becomes identical.

We then define a *Likelihood* $\mathcal{L}_j^F(n)$ for each partition of fibers. Each element $S_{j,k}$ of the partition set is in itself a set of activated fibers, which in our model is by definition an observable O_m. We can look for them inside of the *Likelihood* matrix and multiply all of the resulting *Likelihood* distributions like we did with the toy model in equation 4.8 to produce a combined approximated *Likelihood*.

$$\mathcal{L}_j^F(n) \;=\; \mathcal{N}(K_j)\cdot\prod_{k=1}^{K_j}\mathcal{L}(n,S_{j,k}) \tag{4.12}$$

What we are implicitly assuming here is that the excitation of the fibers has been caused by different multiple particles hitting the detector, when we partition the fibers into K sets we are assuming that the excitation of each of these sets has been caused by a single particle, hence the normalization factor $\mathcal{N}(K)$ that holds little significance here (the normalization of a *Likelihood* is arbitrary) but will be vitally important in the next step. Let's also note that since after a few units $\mathcal{N}(p) = 0$ we need not, in practice, to list all of the partitions of F.

To compute the approximated likelihood of the entire event we average the contribution of all possible partitions similarly to what we did for the toy model in equation 4.9:

$$\mathcal{L}^F(n) \;=\; \frac{1}{2^{N-1}}\sum_j\left(\mathcal{N}(K_j)\cdot\prod_{k=1}^{K_j}\mathcal{L}(n,S_{j,k})\right) \tag{4.13}$$

What we wrote is the likelihood we associated with an exact match, which means that $\mathcal{L}^F(n)$ does not account yet for noisy or inefficient fibers.

To account for these two effects we can list all[9] of the possible subsets of the 796 fibers of the entire FACT detector. Let E be one of these subsets and e_i be its elements. For each of these subsets we can write a different fiber activation scenario in which the fibers that are present in E behaved in a faulty way; namely were $e_i \in F$ we will assume the fiber fired due to noise, were $e_i \notin F$ we will assume the reason why the fiber did not fire is its finite efficiency.

We can write under the E hypothesis which set of fibers scenario would have been activated if the detector were an ideal one, namely:

$$F \triangle E \;=\; (\, F \cap E^c \,) \;\cup\; (\, F^c \cap E \,) \tag{4.14}$$

where the apex "c" indicates the complementary set (with respect to the entire FACT detector). In this scenario we removed $r = |\, F \cap E \,|$ fibers marking them as noisy and added $a = |\, F^c \cap E \,|$ marking them as inefficient. Given the probability η of a fiber hit by a particle to not be activated[10] and ϵ the probability of a fiber that was not involved in the decay to be nonetheless activated in the coincidence time window. We can write the probability $\mathcal{P}_E$ associated with the fault scenario E as:

$$\mathcal{P}_E(\eta, \epsilon) \;=\; \epsilon^r \cdot (1-\epsilon)^{N-r} \cdot \eta^a \cdot (1-\eta)^{796-N-a} \tag{4.15}$$

It can be noted that the probability $\mathcal{P}_E$ is correctly normalized, i.e.:

$$\sum_E \mathcal{P}_E \;=\; 1 \tag{4.16}$$

To compute the likelihood relative to the noisy detector model we need to do nothing more than to sum over all of the possible choices of E the associated *Likelihood* $F^*_E(n)$ weighted by the scenario probability $\mathcal{P}_E$:

$$\mathcal{L}^*(n)(\eta, \epsilon) \;=\; \sum_E \mathcal{P}_E(\eta, \epsilon) \cdot \mathcal{L}^{F \triangle E}(n) \tag{4.17}$$

Since as stated in equation 4.15 $\mathcal{P}_E$ is a polynomial in η and ϵ, $\mathcal{L}^*(n)(\eta, \epsilon)$ will be a polynomial too. Due to ϵ being close to zero, higher terms in ϵ of the polynomial will have an exponentially depressed contribution (in practice we do not expect any term above ϵ^2 to be of any significance. The parameter η

[9] Although in practice we will have to weasel ourselves out or listing all of these instances if we want the reconstruction to be computationally feasible.

[10] That in this formulation we are approximating as a constant.

might indeed be less than 0.5, which would mean that $\mathcal{P}_E$ would be higher for terms in which E activates more fibers overall. Still as we turn on more and more fibers the only nonzero terms in the sum over the partitions of equation 4.13 will be those in which the partition assumes that many particles were emitted from the same vertex. As we said before the probability $\mathcal{N}(q)$ of emitting q particles either decreases exponentially or abruptly becomes zero.

Therefore only a few terms of the potentially enormous polynomial expressing $\mathcal{L}^*(n)(\eta, \epsilon)$ will actually be relevant, in practice a score of terms will suffice for any real life application. The computational cost of calculating the coefficients of such terms (coefficients that, let us note, are *Likelihoods*) is potentially enormous. Most of the sums and products that we wrote to reach this result (equations 4.13 and 4.17) expand into trees most branches of which have null or negligible contribution. To the best of our ability to give a ballpark estimation, based on similar problems we tackled in the past, we do not expect the number of meaningful terms in the worst case scenario to exceed 10^9, which means that if the correct pruning techniques are employed, it can be done. Of course the technical challenge posed by this calculation is not at all trivial and much care is necessary when designing the tree enumeration schemes and the proper containers to store lookup tables and precomputed *Likelihoods*.

4.7 Preliminary results

As stated before, the FACT detector has never been designed to reconstruct the r coordinate of vertices. The detector is currently being used successfully to reconstruct z positions, burst intensities and timings and in general to diagnose the experiments. Vertex reconstruction[11] based on Bayes regression is a future upgrade to the traditional vertex reconstruction in the detector and is currently still being developed and tested.

As of today we developed the geometry-only Monte Carlo simulation code and optimized it aggressively. We wrote the necessary code to compute the *Likelihood* matrices for single and two-particle decays in FACT and the basic reconstruction code. We solved the problem of finding a way to effectively store the *Likelihood* matrix and fulfill all of the requirements of an efficient implementation of the algorithms enunciated so far, by splitting the matrix into columns, dividing them by number of active fibers and saving them onto disk as sorted vectors in which the sorting is given by interpreting the status of the FACT fibers as a single 796-bit-long integer and employing the natural integer order. We also developed utilities to manage, lookup and update entries in this kind of container that we employ to run our tests. We

[11] Or, as should be now clear to the reader, we should better say: "discrimination between hypotheses on the vertex distribution"

are currently still developing the necessary code to compute the complete combinatoric likelihoods.

By testing the reconstruction against events generated with the same simulator used to build the *Likelihood* matrix we were able to determine the $r - z$ position of point sources placed inside of the detector with an accuracy of less than 1 mm using a few hundred events (see figure 4.11). We saw no significant difference between employing one-particle or uncorrelated two-particle decays (we keep employing the same model to build the matrix and to test it), but for the factor two given by the decay product count.

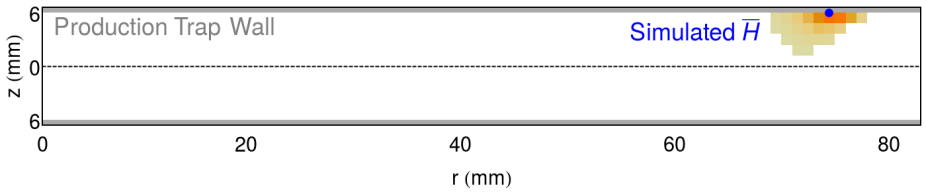


Figure 4.11: Combined likelihood for 100 simulated $\bar{\mathrm{H}}$ annihilating on the production trap wall, each emitting a single pion during its decay. The likelihood was computed using directly the $\mathcal{L}$ matrix computed for a single event.

We also attempted the reconstruction of events using only part of the *Likelihood* matrix. Usually in our test we let the matrix blow up to 10^8 observables after which we stop adding new columns to the matrix and proceed only to collect more statistics for the columns we already hold. We saw that if, after building the matrix, we restrict ourselves to employ only the 10^7 most common observables in the two-particle *Likelihood* matrix we are still able to reconstruct events.

We are currently developing the code needed to compute multi-particle *Likelihoods* from single and two-particle *Likelihoods*, the main challenge of the task being to find the optimal way to enumerate the combinatorics required by formulae 4.13 and 4.17. We are currently gauging the effectiveness of several enumeration orders before we settle for the final implementation.

Conclusions

He could have avoided a lot of anguish if he had just followed the data to their obvious conclusion and admitted the stars probably didn't exist.

– Alexander Scott
The Unsong Book

Decades of tests on the different ways matter and antimatter interact have failed in revealing the presence and magnitude of any gravitational interaction of antimatter. A possible probe to achieve this result is offered by $\bar{\mathrm{H}}$. The AE$\bar{\mathrm{g}}$IS experimental setup is on the verge of starting the production of $\bar{\mathrm{H}}$ through the charge exchange mechanism.

Successful production of $\bar{\mathrm{H}}$ at AE$\bar{\mathrm{g}}$IS will need to be confirmed and a possible path to do so is potentially provided by the Bayesian vertexing algorithm that we developed for the FACT detector. In this work we have shown the theoretical validity of such technique through the use of toy models. We have furthermore shown the feasibility of such reconstruction technique in the ideal case of a noiseless detector, whereas the noise-resistant version of the algorithm is currently being developed.

To perform $\bar{\mathrm{H}}$ production through charge exchange a large quantity of cold positronium excited in Rydberg states is required. We have demonstrated the ability to produce Rydberg-excites positronium in the $\bar{\mathrm{H}}$ production zone of AE$\bar{\mathrm{g}}$IS, located inside of the superconducting 1 T magnet. This result has been achieved due to us developing proper diagnostic tools and a novel data analysis technique for the SSPALS spectra capable of detecting faint signals in a slowly drifting environment.

The development of these diagnostic and analysis tools has been possible thanks to the commissioning of a specialized equipment: the *Breadbox*. The *Breadbox* allowed us to study the Rydberg band excitation of positronium to a level of detail that could not have been achieved inside of the 1 T magnet. Moreover the possibility to operate inside of this privileged setup allowed

us to achieve two novel Ps spectroscopy results: excitation of the 3^3P level of positronium from the 1^3S state and production of the metastable 2^3S positronium.

Production of cold positronium in AE$\bar{g}$IS is always mediated by nanochanneled silicon converters (NCPs). Since an ever increasing availability of cold positronium is crucial to our ongoing endeavors we tackled the problem of identifying which converters might offer better future performances.

We developed a novel simulation code capable of handling a geometry that mimics the complex nanoscopic structure of the NCP converters. Through the use of such code we investigated the implantation and diffusion of e^+ in NCPs, and identified energy ranges in which these processes can be described through the use of simpler models such as Makhovian profiles or Monte Carlo simulations in bulk materials.

We then developed and calibrated a classical model for the transport and cooling of positronium atoms within the nanoscopic channels capable of:

- Correctly describing the behavior of NCPs held at room temperature as described in the current literature.
- Correctly predicting and explaining why the energy spectrum of the positronium emitted by NCPs is shaped like the superimposition of two thermal distributions, fact that had been observed but thus far had eluded explanation.

Direct application of this result is, in the case of NCPs held at room temperature, the capability of predicting the performance of NCP converters and optimize them before they are built. This capability we are already applying to our latest endeavors.

Future application of this result is the possibility, at least theoretically, to predict the performance of converters that operate in the transmission scheme instead of the reflection one. This is of great interest since the amenability of such converters would greatly simplify the design of the Ps experimental setups.

Through ToF measurement we determined the energy spectrum of positronium emitted from NCP converters to a higher degree of precision, and for lower sample temperatures, than what was done in the past. We confirmed the presence of a significant fraction of positronium with a kinetic energy comparable to that of the constituents of NCP converters held at cryogenic temperatures. This fact has two fundamental consequences:

- It shows that is possible to employ NCP converters in experiments that require Ps at cryogenic temperature.
- The model that so effectively describes NCP held at room temperature does not apply to NCPs held at cryogenic temperatures. In particular

from the literature we can see that the model holds at least down to 150 K while new measurements show it breaking down at 20 K.

Our interpretation of the latter consequence is the transition, during the cooling process that the positronium undergoes in the NCP channels, to a regime in which the quantum nature of the Ps atom becomes relevant; phenomenon, this, that had been theorized in the past but never observed directly. This suggests that attempting to further thin the NCP channels is not likely to produce colder positronium, whereas designs such as conically-shaped channels or multiple stages of cooling with progressively sized channels are more likely to provide a valid successor of our current NCP converters.

Acknowledgments

It is a known custom that in a thesis, one takes some time to acknowledge every individual who has contributed, either materially or spiritually, to the research work which has culminated in this final edition. Yet , if I were to list all those deserving to be thanked this section would extend beyond what a sensible limit should be. So let me begin by extending my gratitude to all those who lent their hands, knowledge, wisdom, encouragement and, that most precious of resources, time to me over the last three and a half years. However I'd be remiss to not indulge in some notable mentions, whose contribution has been significant if not essential to the works' successful completion.

First of all **Davide Orsucci**, who had the misfortune of becoming my husband last year and who is the invisible catalyst behind my career. Ever so often, when I ask myself on how I came to understand something, I cannot but admit that he was the one who was able to explain it to me. Moreover his unbreakable blind faith in my ability to overcome any challenge has been, and continues to be, irreplaceable motivational fuel.

Working in Trento in the last three years meant I had the privilege to collaborate with a wonderful team. **Luca Ravelli**, who figured out that selling wine is more profitable than manipulating positrons. **Luca Penasa** to whose electronic creations you could blindly entrust your own life. **Marco Bettonte**, who knows of the unsung beauties of the machine workshop and **Sebastiano Mariazzi**, living embodiment of the chaotic good alignment.

Then there is the AE$\bar{g}$IS collaboration that welcomed me in their collective endeavor and then provided guidance, knowledge and challenges. Within the collaboration, **Ruggero Caravita** following whose teachings I might one day become a successful scientist, or get thrown in jail, or both. **Stefan Haider** and **Chloe Malbrunot** whose encouraging feedback let me endure through that special kind of frustration that only uncooperative machinery can elicit. **Giancarlo Nebbia** who had me play with the most expensive electronic component I ever held in my hands. **Germano Bonomi** who trusted me the keys to a project he is heavily invested in. **Francesco Prelz** who showed me how high the power ladder of informatics can go. **Gemma Testera**, **Daniel Krasnický** and the entire Genoese team, my first interaction with which was to accidentally dismantle part of their apparatus. **Benjamin**

Rienaecker, **Milena Vujanovic**, **Lillian Smestad** and **Philip Hackstock** who, by playing with my programs, fulfilled my dream of becoming a toy maker.

Furthermore **Werner Egger** who led me inside of a nuclear reactor's hall and **Neofytos Themistokleous** who struggled with the implementation of colossal Monte Carlo simulations.

Then **Alessandro Moia** with whom I came to the agreement that reality does not exist. **Nadim Zein** who would probably warn me that reality is an illusion created by language. **Chiara Strazzulla** who might be uncertain on the existence of reality but would bet that if there were any way to destroy it, physicists would find it. **Andrea Sechi** who would attribute any discourse about destroying reality to unconfessed control manias in my subconscious. **Attilio Nizzero** who would advise me against getting too narcissistic about my mental trips about the existence of reality. **Valentina Servile** who gifted me the wonderful year in which she let me be her house cat. Wonderful because cats need not to worry about the existence of reality as long as someone is scratching their back.

My parents that periodically showered me with their unquestioned confidence about my potential. And took care of **Maxwell** while I was in Trento.

Finally **Roberto Brusa** who phoned me when I was looking for a PhD "anywhere but Italy" and instead introduced me to a great Italian research team. Then trusted me his equipment, allowed and financed my training and my own secondary projects and in general gave me maneuver space while at the same time reining me in whenever my enthusiasm dragged me away from the final goal. He showed me how great scientific endeavors should be designed as long paths of relevant incremental achievements and how unpublished results are of no benefit to the community. He showed me many ways in which even smart people can lose themselves and achieve no scientific results and he continuously advised me about not putting my face in the trajectory of someone else's swinging fist. He advised me to read more literature and to avoid reinventing the wheel. Two years ago he got me stunned when he asked a question that which might be the most important of my life; question to which as of today I still have no answer. As I approach the end of my PhD, the feeling that I get from the thought that I'll have to walk by relying more on my own legs and less on his guidance is almost scary.

Thanks.

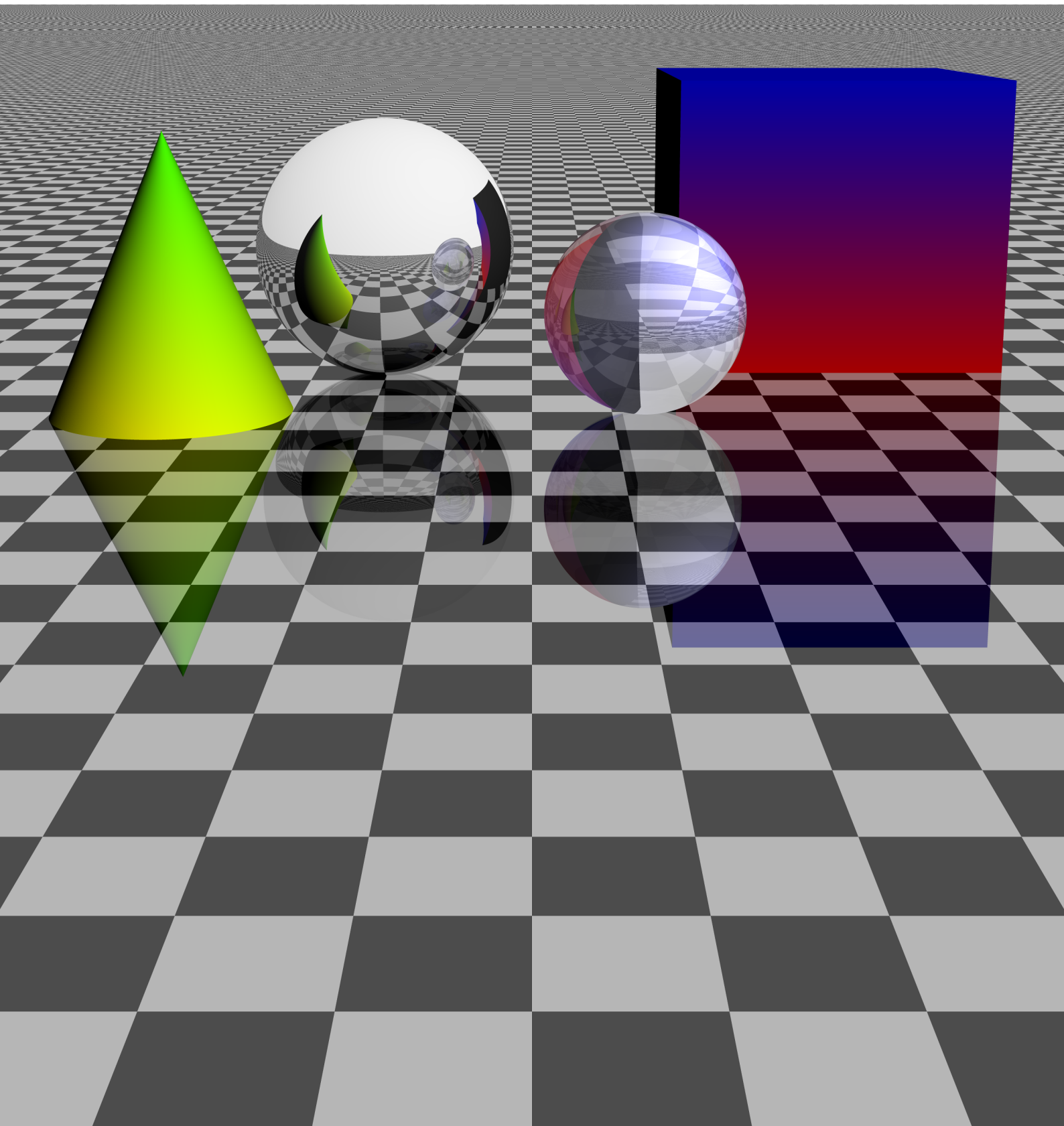
Abbreviations

ACAR – Angular Correlation of Annihilation Radiation
AD – Antiproton Decelerator
ADC – Analog to Digital converter
API – Application Programming Interface
BEC – Bose-Einstein Condensate
CAD – Computer Assisted Design
CCD – Charge-Coupled Device
CF*xx* – xx mm-diameter ConFlat flange.
CLI – Command Line Interface
CMOS – Complementary-MOS
CPT – Charge, Parity and Time reversal
DAQ – Data AcQuisition
DCAM – Digital CAMera protocol
DLL – Dynamic-Link Library
DMA – Direct Memory Access
FACT – Fast Annihilation Cryogenic Tracker
FWHM – Full Width Half Maximum
FPGA – Field Programmable Gate Array
GCC – Gnu Compiler Collection
GPIO – General Purpose Input/Output
GUI – Graphical User Interface
$\bar{\mathrm{H}}$ – anti-Hydrogen
HTTPs – Hyper-Text Transfer Protocol over Secure socket layer
HV – High Voltage
IC – Integrated Circuit
IIDC – Instrumentation & Industrial Digital Camera consortium
LEAR – Low Energy Antiproton Ring

MCA – Multi Channel Analyzer
MCP – Micro Channel Plate
MCU – Micro Controller Unit
MPPC – Multi-Pixel Photon Counters
NaI(Tl) – Thallium-doped crystalline sodium iodide scintillator
NCP – NanoChanneled Plate
$\bar{\mathbf{p}}$ – anti-proton
PDF – Probability Distribution Function
PET – Positron Emission Tomography
PMT – Photomultiplier Tube
PHP – PHP Hypertext Preprocessor
PRINCEPS – Passage / Reflection In Nanochannels, Convertible Experimental Positronium Source
Ps – Positronium
PSP – Porous Surface Plate
PSU – Power Supply Unit
QCD – Quantum ChromoDynamics
QED – Quod Erat Demonstrandum – What was to be proven
QED – (alternatively) Quantum ElectroDynamics
QFT – Quantum Field Theory
RF – Radio Frequency
RS232 – Electronic Industries Alliance Recommended Standard 232 serial bus
SEM – Scanning Electron Microscope
SSH – Secure Shell protocol
SSPALS – Single Shot Positron Annihilation Lifetime Spectroscopy
ToF – Time of Flight
TTL – Transistor Type Logic
URI – Unique Resource Identifier
USB – Universal Serial Bus
VCR – Video Cassette Recorder / Video Cassette Recording
V-I – Voltage-Current response of a two terminal component
VISA – Virtual Instrument Software Architecture

Appendices

3D rendering of simple geometric primitives obtained by running the same codebase employed to simulate the behavior of positrons inside of nanochanneled plates. To produce a rendering the portion of the code that implements the interactions of positrons with matter are replaced with an implementation of a simple approximation of the physics of visible light.



Appendix A

Tracing Complex Geometries

The mind game is a relationship between the child and the computer. Together they create stories.

– Orson Scott Card
Ender's Game

As said in Chapter 1 to be able to simulate the behavior of particles inside an object of a given geometry we need to be able to compute two functions of that geometry: the *InclusionCondition* and the *NearestIntersection*. The first is defined as a function that, given any point in the three dimensional space, determines if such point is contained within the volume of the solid object or not. The latter is defined as a function that, given a half line in the three dimensional space in the form of an origin and a direction, returns, if it exists, the nearest intersection between the half line and the surface of the solid and the normal to the surface on the intersection point.

A slight correction should be mentioned before proceeding further. We can slightly modify the *NearestIntersection* condition so that alongside the origin and direction of the half line also a maximum relevant distance d_M is specified. That is: if the distance from the origin of the nearest surface intersection d_s is greater than d_M then we don't care about obtaining the exact solution and the code is allowed (but not required) to return $d_s = \infty$. This is useful since ranged *NearestIntersection* problems allow more and more powerful optimizations than unbounded ones. Every time we trace the path of a particle inside of a bulk we can determine from the interaction model a randomly generated interaction distance d_i before checking for surface intersections; since the bulk interaction will take precedence over the surface collision whenever $d_i < d_s$ we can formulate bounded *NearestIntersection* problems and benefit from the optimizations.

We implemented code that solves both problems for a set of primitives; of interest for the simulation in nanochanneled plates are cuboid surfaces, spheres, spherical slices and truncated cones. In all of these cases the *InclusionCondition* is trivial, whereas the *NearestIntersection* tends to require slightly more computation. For quadratic surfaces the nearest intersection problems can be reduced to computing the solutions of a quadratic equation, task that needs to be undertaken carefully by first computing the greatest solution from the resolving formula and then obtaining the other solution from the first one and the solution product. Attempting to compute directly both solutions from the closed formula is likely result in glitches deriving from catastrophic cancellation.

A.1 Raytracing rendering

One of the main hardship of implementing simulations in complex geometries is the necessity to debug thoroughly a code that might glitch in unexpected ways and on rare occasions. Several techniques exist that can trap such glitches, such as bulk checking after surface traversal. A good practice is to embed them in the simulation code, enclosed in preprocessor directives that make them disableable en masse at production.

Another good way to debug the geometric code is to have it produce a graphic representation of the described geometry through the process of rendering. Due to the reversibility of optical paths we can, in fact, use the *NearestIntersection* condition to approximate the color of the light collected by each pixel the CCD of a virtual camera used to photograph the scene, thus extracting a graphical representation of the scene.

The procedure works as follows: first of all a global illumination is chosen and a color is assigned to each primitive present in the scene. Then a point is chosen in space as the ray casting origin (more informally, the point-of-view from which the geometry is observed) and a rectangle with the aspect ratio of the image to be produced is placed in the three dimensional space so that the ray casting origin lies in its axis. The rectangle's axis is the direction from which the scene is observed. Then the rectangle is subdivided into a grid of as many square pixels as desired and for each of them a half line is traced from the origin to the center of the pixel. The object which yields the closest *NearestIntersection* will determine the color of the pixel, the angle between the surface normal on the impact point and the direction of a chosen global illumination the pixel's luminosity, see an example of this in figure A.1a. The technique can be extended to simulate partially reflective or refractive materials by tracing secondary rays from the impact points in the same way that secondary particles are traced in Monte Carlo simulations of showers. Blending together the colors resulting from the entire trace tree

yields a much more correct estimation of the apparent color of the pixel. An example of this, obtained through the same code employed to run the implantation/thermalisation simulations can be seen in figure A.1b. If this figure might hint to the more general technique of realistic 3D rendering it is not by accident. Although what I mentioned here is just the tip of the iceberg, the art of simulating view and illumination is a complex discipline that, although implementing the physics of a completely different energy range, shares the same roots as that of simulating interactions between radiation and matter and widely employs similar techniques.

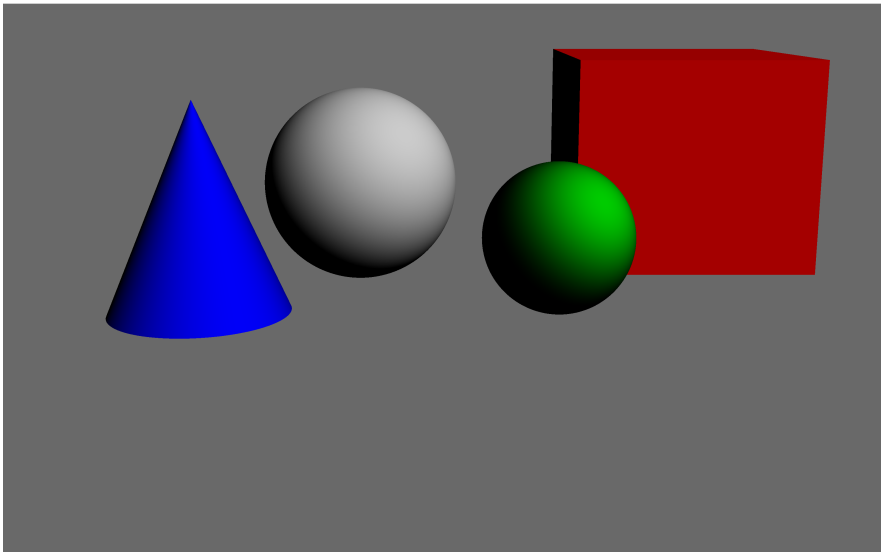


Figure A.1: Above a simple example of rendering by raytrace employing the primitives implemented to run the simulations presented in Chapter 1, the same rendering when a more advanced model of interaction is employed and the traced photons spawn secondary reflected/refracted photons from the points of impact on surfaces can be see in the opening graphics at the beginning of this appendix.

A.2 Clusters and Neighborhoods

Let's now take a closer look at how we transition from solving the *Inclusion-Condition* and *NearestIntersection* for truncated cones and spheric slices to solving them for dendrites and NCPs.

First of all to transition from the *InclusionCondition* for a single solid to the *InclusionCondition* for a collection of them we need nothing more than applying an or condition. To switch from the *NearestIntersection* for primitives to the *NearestIntersection* for union of solids we need to enlist all of the collisions with all of the primitives, then to exclude those that lie inside of a different primitive than the one that produced the intersection. Finally we accept the one intersection closest to the half line's origin point.

This is true of single dendrites but can also be applied to arbitrarily large dendrite collections. As said in chapter 1 we subdivide the NCP macroscopic surface into cubic *Voxels* having a side that exceeds the maximum length allowable for a dendrite in any given direction. Therefore any intersection that takes place inside a specific voxel has to happen on the surface of a dendrite whose aperture is located in the upper face of the either specific *Voxel* or one of the adjacent ones (see figure 1.3b). Therefore in general we can compute NCP intersections by first computing in which *Voxel* the origin of the half line is situated, then by computing the *Cluster* relative to that specific *Voxel* defined as the set of all of the dendrites whose apertures are collocated in the *Voxel's* upward face. After computing the *Clusters* relative also to the eight adjacent *Voxels* we can finaly combine the *NearaestIntersections* of the nine adjacent *Cluster* and, in case the line propagates over the central *Voxel's* edge, reformulate the problem from the intersection with the voxel edge onward.

This method is highly inefficient for two reasons. The first is that when computing *NearaestIntersections* with a solid composed of the union of primitives it is in general more efficient to compute the union of n primitives altogether instead of computing union of unions of primitives due to several shortcuts that are possible in the latter case. Second reason why this method is inefficient is that most dendrites originating from *Voxels* adjacent to the analyzed one do not intersect the analyzed *Voxel*. To avoid these inefficiencies we then define the *Neighborhood* as the collection of all of the geometric primitives (truncated cones and spherical slices) belonging to dendrites originating from nine adjacent *Clusters* that intersect the central *Voxel*. Of course to generate a single *Neighborhood* it is necessary to generate all of the adjacent *Clusters*, as we will see due to caching this does not constitute a problem. *Neighborhoods* are what we then, actually, employ to solve *InclusionCondition* and *NearestIntersection* problems in the NCP.

A.3 Optimizing raytracing

The greatest challenge we face when we try to perform radiation matter interaction simulations in complex geometries is that of minimizing the time required to solve the *InclusionCondition* and *NearestIntersection*. The amount of primitives that constitute a single *Neighborhood* geometry can become

extremely large and computing the result of *InclusionConditions* and *NearestIntersections* for all of them is a prohibitively inefficient way to solve the two problems for the entire geometry.

A first optimization is to compute bounding boxes for all of the employed primitives and to check the bounding boxes for intersections before checking the underlying geometry. Naturally this optimization will yield benefits only if the probability of the half line intersecting the bounding box is less than:

$$1 \; - \; \frac{t_b}{t_p} \tag{A.1}$$

with t_p the time needed to compute the intersection with a primitive and t_b the time needed to compute the intersection with a bounding box. When I benchmark the code of our simulator on my portable processor[1] I find an average $t_p = 579\,\text{ns}$ and $t_b = 93\,\text{ns}$, so as long as at least 16% of the potential collisions are excluded a priori by bounding box checking the optimization will produce a benefit.

Another possibility to implement bounding box optimization is to implement an imperfect discrimination, that is instead of attempting to determine precisely if the half line intersects the bounding box or not we aim only at excluding the most obvious situations in which we can determine that no intersection takes place. If $\vec{o}$ is the half line origin, $\vec{d}$ its direction in space and $\vec{v}^{min}$, $\vec{v}^{max}$ respectively the cuboid vertexes with the smallest and greatest coordinates, the condition:

$$\left(o_x < v_x^{min} \;\wedge\; d_x < 0 \right) \;\vee\; \left(o_x > v_x^{max} \;\wedge\; d_x > 0 \right) \tag{A.2}$$

detects only bounding boxes that do not intersect the half line; the same condition can be duplicated for all three dimensions to obtain a fast bounding box discriminator which runs in about 32 ns

This optimization is not nearly enough to allow us to perform the simulation without spending a great deal of computational power. A typical *Neighborhood* in our NCP model contains about $0.5 \cdot 10^6$ geometric primitives. Even if all of them were to be excluded by bounding box checking each interaction would require 16 ms which would result in about three minutes to compute the implantation of a single positron and into potentially hours to compute the thermalisation of a single positronium inside of the channel.

We need to add a pseudo-octree optimization. The main idea is that of collecting bounding boxes that are spatially close and to compute their overall bounding box. Then to aggregate again the resulting bounding boxes up until

[1] Bearing in mind all of the complexities associated with benchmarking on a modern processor, these numbers should be taken with a grain of salt

all of the primitives are included in a single grand total bounding box, while in the process we built a tree-like structure of bounding boxes in which boxes in each level are entirely contained in the parent box. Then to determine the relevant primitives we just need to check if the parent bounding box intersects the half line and descent into its child boxes only if an intersection is found. An example of this technique in action is shown in figure A.2. The effectiveness of this technique is highly dependent on the algorithm employed to aggregate the bounding boxes to build the box hierarchy. In our case about a 10^3 bounding box evaluations are required to identify about 10^2 relevant primitives, this translates into about 150 µs per particle interaction and 1.5 s per implanted particle on a single core, values that are compatible with the actual performance of the whole simulator.

Finally a word needs to be spent on the algorithm employed to build the tree. Computing the optimal tree is an extremely computationally-heavy process that would defeat entirely the optimization purpose. Instead we tested several heuristic algorithms to achieve a good balance between tree construction time and tree performance. We settled onto the following one.

We build the tree concurrently to the *Neighborhood* generation. The structure we employ to build the tree consists of nodes that contain an aggregate bounding box of their content, a collection of bounding boxes and up to eight children nodes. We start with an empty tree containing only a root node consisting of a degenerate zero-volume bounding box, no children nodes and an empty collection of bounding boxes. As soon as the geometric primitives are added to the *Neighborhood* we insert their bounding boxes to the tree. We add each new box to the root node, then we adopt the following rules for adding boxes:

- If a node has no children then we add the box to the node's box collection and the node's bounding box is extended to comprehend the new box within its volume.
- If a node has children then every box in its collection must be deleted from the collection and added to one of its children nodes. The parent bounding box is nonetheless extended to comprehend the new box within its volume.
- If a node's box collection contains more than 16 elements then the node acquires 8 children nodes and the collection elements are moved from the parent node to the children nodes.

The rule for selecting in which children node a box should be added is the following. The current parent node's bounding box is subdivided into eight octants obtained by splitting in half each of its dimensions and a children node is assigned to each octant. Then the center point of the bounding box

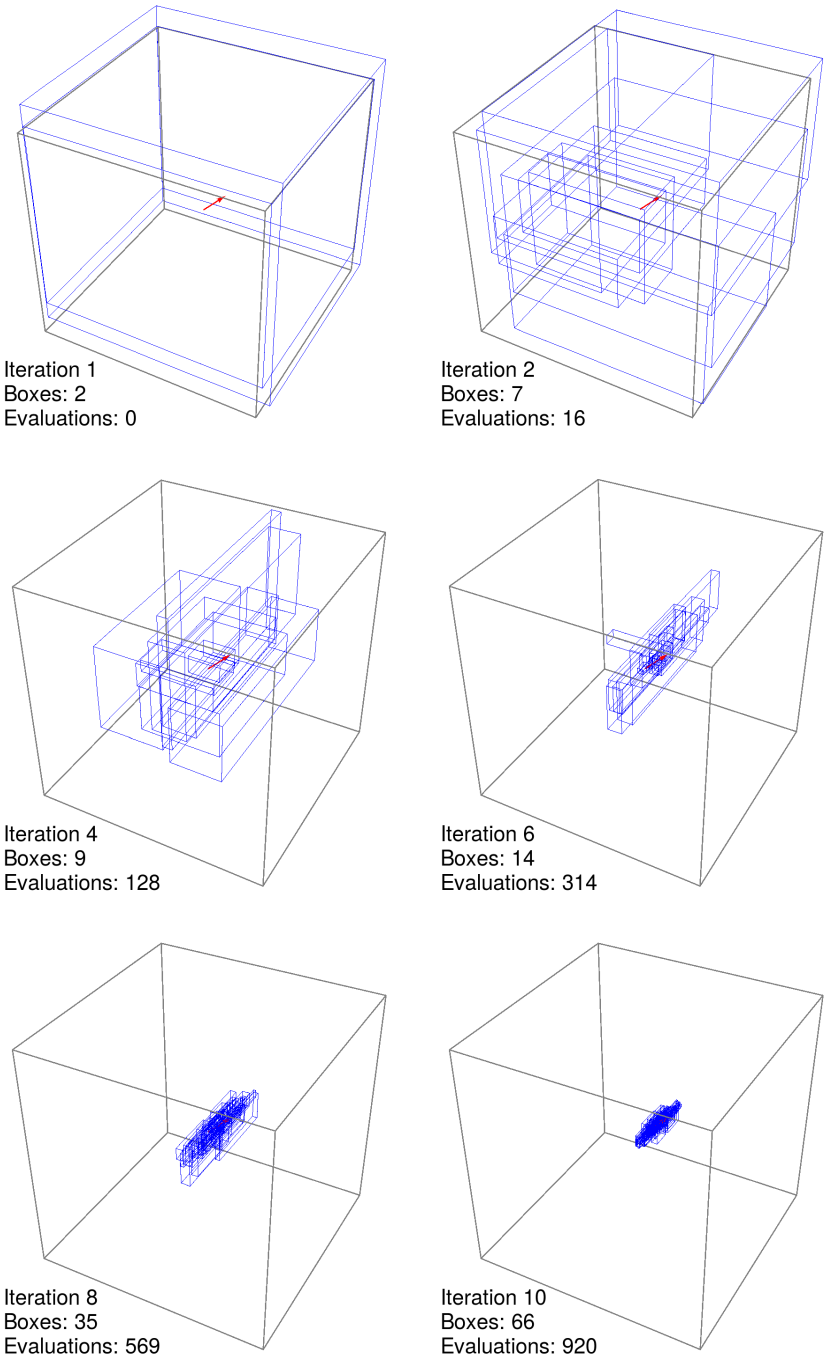


Figure A.2: The hierarchic bounding box technique shown in action as it detects the relevant primitives inside a dendrite *Neighborhood* consisting of 512143 primitives. The entire evaluation can be performed in less than 100 µs

to be added is computed and the box is added to the node corresponding to the octant the central point lies in.

This brief exposition does not cover the entirety of the acceleration techniques we employed to efficiently solve the *NearestIntersection* problem, but should give a general idea of the methods.

A.4 Caching and Multithreading

Reached as we have this point of the description of the computational techniques employed to simulate positron implantation in NCPs an elephant still sits in the room. If we were to run the random number generator to build the dendrite neighbourhood every time we need to propagate a particle from one interaction site to the next we would be wasting most of our computational capabilities into generating dendrites. It is therefore natural to keep a cache of the most recently used dendrite neighbourhoods and use them as many times as possible before recycle the memory space dedicated to storing them and generate different ones.

In our simulator two similarly organized caches are kept, the first one to store recently employed dendrite *Clusters* the other to store recently employed dendrite *Neighborhoods*. The external interface of the caches is fairly simple in principle: the tracing process queries the cache for a specific *Cluster* or *Neighborhood* using the element's URI (in our case the integer (x, y) coordinates of the required *Voxel*), and the cache returns a memory location. Paired with the memory location is an indication of whether the required *Cluster*/*Neighborhood* was found in the cache. If the element was found in the cache the the caller function is free to immediately use the resource found at the memory location returned by the cache, on the contrary if the element was not found in the cache the caller function will have the responsibility of building the resource and store it at the location returned by the cache. When the process has finished using the cache element it has the duty of notifying the cache that the resource can be recycled at will; at the same time the cache grants that until the process has sent a no-longer-in-use notify it will keep the resource available.

The strategy that we employed to implement the cache is that of allocating new resource on the host computer until it has reached a fixed maximum element count, then to always recycle the oldest resource not in use (in which the age of a resource is given by the number of queries ran from the last time that resource was requested). To obtain an efficient implementation we employed a heap to store cache handlers, sorted using the element's URI. Elements in the cache are also stored in a doubly linked queue that stores the URIs of all of the elements in the cache. Each handler in the heap is paired with a pointer to its corresponding element in the queue. To maintain the

correct sorting of the queue, every time an object is queried in the cache it is first looked up in the heap (which requires $O(\log(n))$ operations with n the cache size) and its link in the queue is moved to the head of the queue, operation that can be performed in constant time. It can be shown that the structure can be maintained through operations of insertion and deletion with a $O(\log(n))$ cost per operation which makes maintaining even a fairly large cache almost transparent to the tracing process.

What said up until now holds for single thread operations. Nonetheless if we want to be able to harness the capabilities of a modern processor or, even better, a computation cluster we need to be able to run the computation over several threads in parallel. This requires a proper management of the conflicts in resource usage. The Monte Carlo technique is quite prone to parallelization since the repeated sampling of the integrand function can be performed independently by uncoordinated agents and then the results be summed without harm. The simplest way to implement multithreading in our simulator would have been that of having each thread instantiate its own cache. This is undesirable since each thread benefits largely from keeping a cache as large as possible within the employed architecture. The other, much more difficult way, is that of having a common cache shared by all of the threads.

Extending the caching mechanism to make it thread safe is no trivial task. We started by employing mutexes to make each operation performed on the cache atomic, that is no more than one thread is allowed to query the cache at the same time; only when the query operation has been concluded and the thread is using the requested resource another thread will be allowed to interrogate the cache. Then instead of flagging resources as being used or not, for every slot in the cache we need to keep track of how many different threads are making use of it and declare it disposable only when such number reaches zero.

After that we needed to extend the query operation to allow the polling of several URIs at the same time and grant that the requests will be queued consecutively. This means that if not enough slots in the cache are available to fulfill the request; from the moment the first available cache slot is destined to satisfy the request any new slot being freed in the cache will be assigned to this specific request until enough slots have become available to evade the full query. This is necessary since to build a *Neighborhood* nine *Clusters* are required and the calling thread will have to collect all nine before building the *Neighborhood* and releasing the resource. This can result in a deadlock if the cache is not able to provide room to build the ninth cluster before the first one is released and its memory recycled. In the single thread case this is granted as long as the cache can hold at least nine elements. In the multi-thread fashion this can occur in more subtle ways: imagine a situation in which the cache holds at most ten *Clusters*, two concurrent threads are aiming to build

two distant *Neighborhoods* and one of the two processes holds all of the even slots in the cache while the other holds all the odd ones. Querying clusters atomically prevents this kind of deadlock.

The next ingredient to the implementation of a working multithread cache are *promises* and *futures*. A *promise* is a multithread primitive that allows a thread to return asynchronously the result of a computation, the *future* is the object that is given in lieu of the actual result and that provides access to the result once this is available. When a function returns a *future* instead of an immediate result, the caller process receiving it is allowed to freely manipulate the *future* and perform any operation that does not require access to its content. Still the moment the process that received the *future* tries to access its content, either the value has already been computed and the process accesses it immediately or the computation is still ongoing ad the receiving process will freeze until the *promise* has been satisfied and a value is saved into the *future*.

The best way that I found to visualize *promises* and *futures* is to imagine *futures* as boxes with a hole in the bottom, the *promise* being the hole. The caller process receives the box as an answer to its request and, in a de Saint-Exupéry fashion, is told that the box contains a sheep. As long as the receiving process does not open the box, it is allowed to move it, place it and store it as seen fit; while the called process is always allowed to sneak a sheep n the box from the hole at any time. If the caller process opens the box before time it will freeze on the spot until the sheep has been put in place so from the point of view of the caller process everything looks like the box contained a sheep all time long.

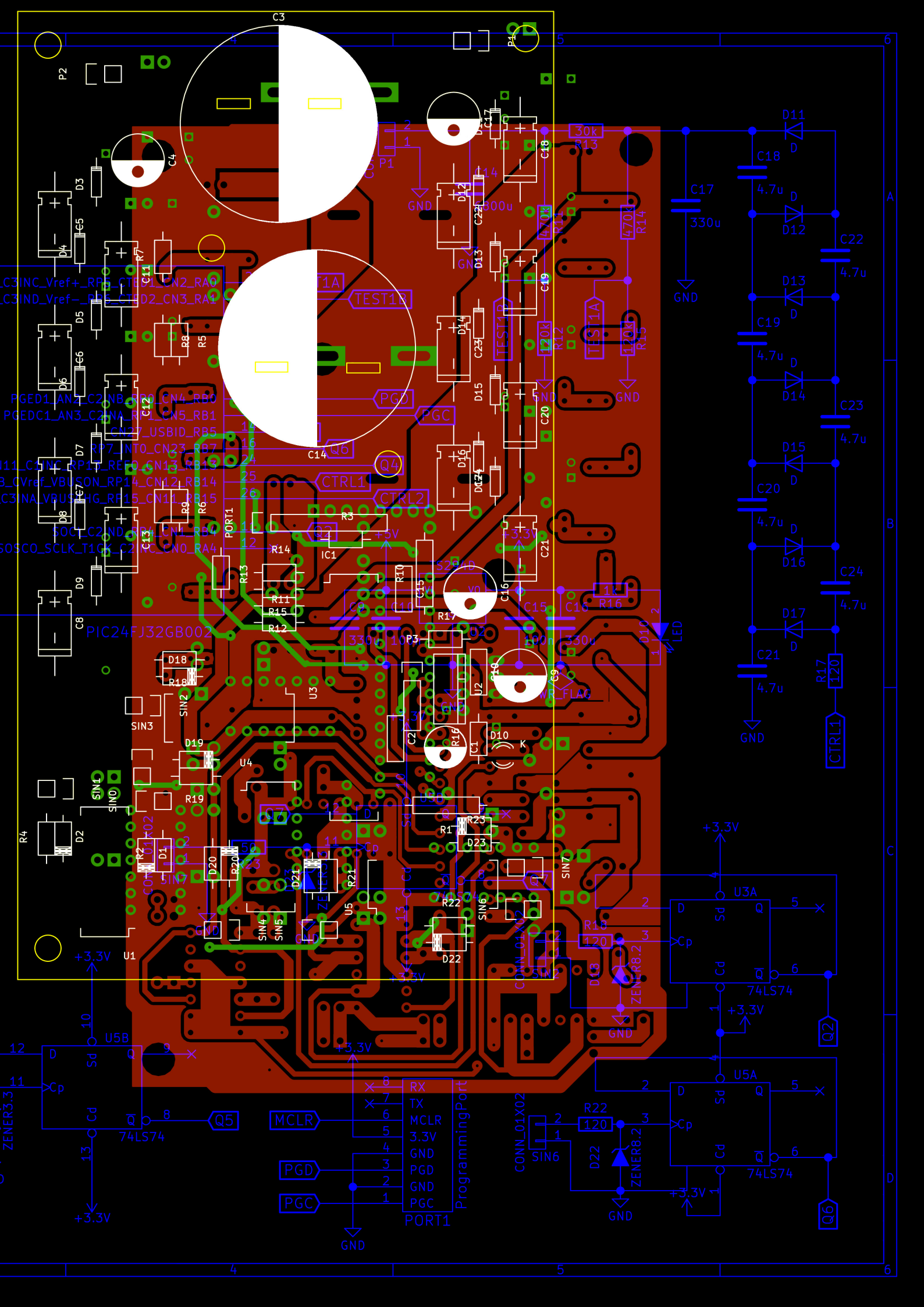
Basic *promises* usually allow only one *future* to be spawned from them, nonetheless it is easy to use them to implement a more powerful object that allows the generation of multiple *futures* from a single *promise*. In our case we need always to make use of *promises* with multiple *futures*, we'll assume that implicitly from now on.

The first function that we need to implement through *promises* is the allocation of slots in the cache. In the single thread application we are always guaranteed that a cache slot can be freed immediately and the query operation can return synchronously an empty slot. In a multithread implementation some time might lapse before a slot can be freed in the cache so instead of being returned a (filled or to be filled) slot in the cache the process is returned a *future* pointing at the slot. If a slot can be made available immediately then the *future*'s *promise* has already been fulfilled, it is appended to a queue of *promises* that are fulfilled as soon as new slots are available.

The second function that we need to refine to allow multithread operation is the avoidance of concurrent usage of partially built *Clusters* and *Neighborhoods*, that is: as long as a *Cluster* or *Neighborhood* is being built no other process should make use of it. To achieve this, in its multithread implementation

the cache does not store *Clusters* and *Neighborhoods* but, instead, it stores multiple-futured *promises* of *Clusters* and *Neighborhoods*. Whenever the cache is queried it will return (wrapped into a *future* as we saw before), a future to the required data and, if the data was not present in the cache, a pointer to the cached promise that the caller thread is required to fulfill the moment it has generated the required *Cluster* or *Neighborhood*. This means that is another process requires the same URI while it is being built, it will freeze on accessing the *future* content until the data has been thoroughly generated.

Detail of the schematics and layers (front and back copper, silk screen and edge cuts) of the fabrication outputs for the PRINCEPS control box.


PIC24FJ32GB002
ProgrammingPort
PORT1
MCLR
PGD
PGC
TEST1A
TEST1B
CTRL1
CTRL2
FLAG
LED
74LS74
CONN_01X02
ZENER8.2
ZENER3.3
+3.3V
GND

Appendix B

Electronics for PRINCEPS

They do as they must do. We may call what they do harmful or useful, but good and evil belong to us, who chose to choose what we do.

– Ursula Le Guin
The Other Wind

During the course of chapter 2 we described how one control operation, namely the setting of the MCP bias voltage, and one diagnostic function, namely the monitoring of MCP and NaI(Tl) rates, have been performed through custom built electronics. We will spend the length of this appendix to detail the design and commissioning of such electronics.

The design requirements for the multiscaler are that it needs to be able to measure pulse rates up to a few kBq with the ability to detect pulses as short as 100 ns. The pulses that are fed to the multiscaler are positive-signed with a variable pulse height ranging from about 5 V to tensions in excess of 12 V. Given this input we opted to assume that the range of incoming pulses that the appliance must be able to handle will need to be prudently $4 \div 30$V.

The MCP voltage is controlled through the analog interfaces of two separate high voltage (HV) power supplies (PSUs). The power supplies provide as output a voltage between zero and 3.5 kV proportional to an input voltage comprised between zero and 10 V. The average input impedance of the HV PSUs across the input voltage range is 250 kΩ. The design requirement for the MCP voltage control is to be able to set it to any value in the $0 \div 2.6$kV range with an accuracy of at least 5 V. On top of that the system must refuse to carry out any request of lowering the HV output voltage with a slope greater than 3 V/s. We limit the speed at which the MCP voltage can ramp (both when rising and lowering it) to increase the MCP lifetime. The 3 V/s

limitation on the falling ramp is not as strict as the ones we actually enforce in the control software since it must only serve as a last-ditch defense in case of a failed control by Melchior, the duty to enforce the actual limitation lies upon the control software. This decision was taken to ensure the possibility of the software to perform an emergency shut down of the MCPs in presence of clear indications of catastrophic discharge inside of them; in such occasion we deemed exceeding what we consider the maximum recommended voltage lowering slope safer than keeping the MCPs powered for a longer time. If possible this restriction should be enforced also in the circumstance of a sudden power loss in the control electronics.

The device we designed was built around a 16-bit Micro Controller Unit (MCU), the PIC24FJ32GB002, chosen out of familiarity due to it having been used by us in previous projects, and because it comprises a built in USB module which provides asynchronous access to the USB interface; having a module that drives the USB port pins autonomously relieves us from having to bit bang them through the DMA which allows us to write a much less tense firmware.

During the design operation we have to reconcile the presence of a variety of different operating voltages across the circuit: the USB interface operates at 5 Vcc, the MCU can drive the USB interface directly but requires a power supply of 3.3 Vcc, the MCP outputs must operate across the $0 \div 10$V range. We would like to develop a compact control device that, if possible, is powered completely through the USB connection.

As long as we are willing to dissipate the voltage drop as heat (and we are) providing a 3.3 Vcc supply from a 5 Vcc is fairly simple. To provide MCP control voltages higher the USB 5 Vcc supply we found that a compact elegant solution consists in merging together the voltage raising and regulation functionalities synthesizing a voltage regulated supply from a Cockroft-Waltron[127] chain. To control the voltage output of the chain we control the frequency of the pulses that we input into the chain itself. Such a chain can be driven directly from an MCU pin as we will detail in section B.1, thus reducing dramatically the number of component required by this construction.

To provide an implementation of the USB protocol on the MCU side we started from an open source stack written by Alan Ott (Signal 11 Software). On the host side we wrote driving software based on the *libusb* open-source library. Of the transfer methods defined in the USB standard we employed only the *Control* and the *Bulk* data transfer types. The device initializes by configuring a single bidirectional endpoint (beyond the standard configuration endpoint 0) which is used for communication with the host. The protocol works by having the host and the device exchange fixed sized frames having a fixed internal structure which is different for frames sent by the host to the device and for frames sent by the device to the host. Frames are XOR

checksummed but devoid of data reconstruction capabilities in case of a failing checksum, therefore we designed the protocol to be resistant to missing frame transfers.

To grant reliable operation of the device during the course of several months we made use also of the MCU's watchdog timer (WDT). The WDT is a built-in cycle counter that unless periodically reset by the MCU firmware will take over control and reset the device. We programmed the WDT to reset the device within a 32.8 s time interval and reset it upon receiving valid frames from the host; this ensures that if due to external disturbances the USB connection is lost (which in our apparatus happened once every few months) the device will force the host to repeat the USB enumeration process which in our case is enough to restore the correct functionality.

B.1 Setting the MCP voltage

As anticipated we employed the Cockroft-Waltron mechanism to produce the $0 \div 10$V signal needed to drive the HV PSUs that power the MCPs. The Cockroft-Waltron mechanism is a passive finite-element transmission line that possess the remarkable property of increasing the voltage of the DC component of the signal that is applied to it adding to it the value of the alternate component's peak voltage multiplied by the number of cells in the series.

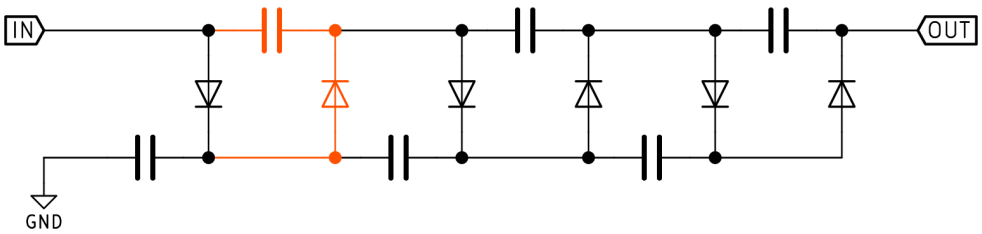


Figure B.1: Discrete element schematic of the Cockroft-Waltron mechanism. Highlighted in orange a single cell.

If we look at how the Cockroft-Waltron mechanism works as a black box, we can describe it (roughly) as a voltage-limited constant current power supply. Let's, for simplicity, restrict ourselves to the case in which the input signal is periodic. The output voltage limit will be given by:

$$V_O = V_p \cdot n + V_{DC} \tag{B.1}$$

where V_{DC} is the 0 Hz component of the driving signal, V_p is the peak voltage

of the alternate component of the input and n is the number of cells in the transmission line chain. The output current I will be given by:

$$I_O \;=\; f \cdot V_p \cdot C \tag{B.2}$$

where C is the capacitance employed to build each link and f is the frequency of the input signal. These quantities hold for the ideal case, in reality the number of total links is limited by the forward voltage drop across the employed diodes and the maximum driving frequency by the capacitor charge profile in a the C-diode circuit. Also formulae B.1 and B.2 hold during a steady state with a fixed output load, whenever the output load is changed about n cycles of the driving pulse must be waited before the system reaches a new steady state.

We dimensioned the Cockroft-Walteron mechanism for our application as a chain of 7 links, with capacitances of 4.7 µF and Schottky diodes with a nominal forward voltage drop of 230 mV. We found that, when driven with a square wave with an output impedance of 120 Ω this chain performs best with input frequencies comprised between 1 kHz and 2 kHz.

The output of the Cockroft-Waltron mechanism is affected by ripples as ample as V_p, therefore to produce a well-conditioned output signal from it, filtering is required. We started by closing the output of the chain directly onto a 330 µF capacitor; since the maximum output current is limited at 2 kHz to 31 mA (which is then further reduced to 27 mA by the expected output impedance of the driving wave) we expect the capacitor to work with the chain's output impedance as a low-pass filter with a cutoff frequency of 2.5 kHz. This means that this filter should in principle clear out all of the high-frequency components of the driving square wave but leave the main driving frequency mostly untouched. In practice due to the Cockroft-Waltron mechanism being highly non Ohmic also the driving frequency gets also canceled out almost completely. From SPICE simulations of the circuitry we expect the voltage ripple after the first stage of filtering to be around 5 mV when the circuit is driven at 1 kHz (half of that at 2 kHz). This is desirable since this is the stage we will be using to sense the output voltage of the Cockroft-Waltron chain and perform a closed loop control of the driving wave.

The voltage obtained after the first stage of filtering is then divided with a 470 kΩ resistor stacked on top of a 120 kΩ resistor. The values ensure that the voltage between the resistors will be clipped in the $0 \div 3.3$V range as long as the voltage at the end of the Cockroft-Waltron chain is comprised in the $0 \div 16$V range. The reduced voltage is then routed to a pin of the MCU equipped with a 10-bit ADC that is used to determine the voltage at the end of the chain. The input impedance of the ADC module is not given by the MCU specifications, still the pin is granted to absorb less than 1 µA, this value is comparable in size with the current that is drawn by the resistive voltage

divider, which means that the ADC measurement will not be linear and that calibration will be necessary to determine a control curve to translate ADC values into output voltages.

The voltage of the first stage is then fed into another RC low-pass filter realized with a $30\,\mathrm{k}\Omega$ resistor and a whooping $6.8\,\mathrm{mF}$ capacitor resulting in a characteristic time of $204\,\mathrm{s}$. This has two effects. The first is that the rippling after the filter is reduced to $5 \cdot 10^{-9}$ ideally (in practice the expected ripple is dominated by the ripple $3.3\,\mathrm{V}$ voltage dropper that generated the MCU power supply, nominally less than $536 \cdot 10^{-6}$). The second is that the maximum slope allowed for the HV output is $17\,\mathrm{V/s}$ upwards and (due to the fact that the diodes in the Cockroft-Waltron chain do not allow the capacitor to be discharged through it) and $2.5\,\mathrm{V/s}$ downwards. This means that even in the case of a disruptive power loss (e.g.: if the USB cable is plugged from the host computer or the computer power-cycles unexpectedly) the MCPs are granted to power down safely as long as the HV PSUs are powered and connected to the controller.

The output voltage of this second stage is divided with calibrated resistors in the same manner the output of the first stage was therefore allowing us to measure the voltage across the controller output. This voltage is the one that is employed by the control software in the host computer to make the necessary control decisions, such as whether requiring a specific voltage will break or not the safety limitations for voltage ramping or if the voltage is within acceptable parameters for the signal coming from the MCP to be considered reliable. Both the voltage measured across the first and the voltage measured across the second stage are sent to the computer for logging and diagnostics.

The complete scheme of one of the MCP control outputs is shown in figure B.2.

B.2 Multiscaler

Measuring the rate of a signal composed of randomly occurring pulses boils down to counting how many of these signals are observed within a given time window. This requires the hardware to be capable to form the incoming signal into a discriminated signal with definite logic levels, then to count the resulting pulses. The pulse counter is then be read out by the control logic and the difference between successive readouts divided by the time difference between the moment these readouts are performed is the measured pulse rate. Since no counter features an infinite amount of digits the control logic needs to read out the counter often enough so that the probability of the counter overflowing between successive readouts is negligible when compared to the required accuracy of the instrument.

The chosen MCU offers 19 General Purpose I/O (GPIO) pins. Two of these

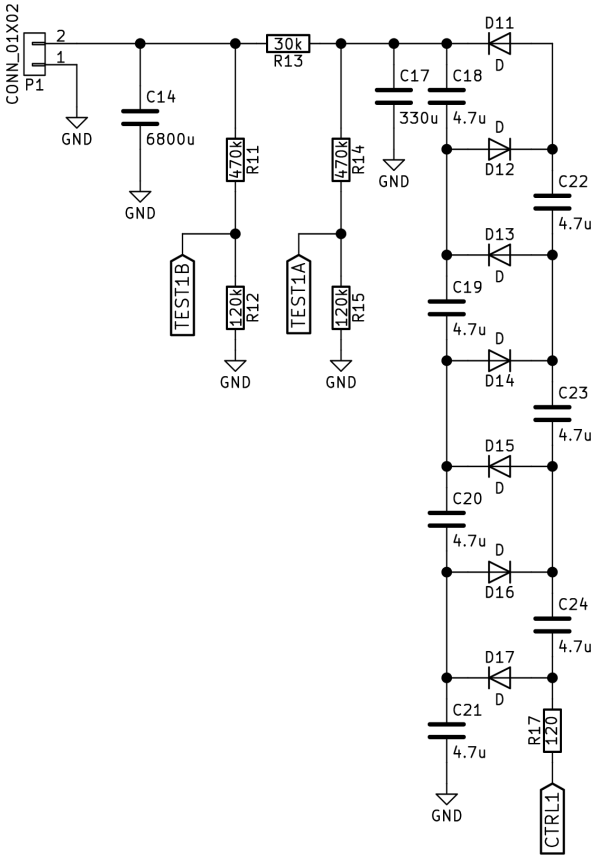


Figure B.2: The schematic of an MCP control line going from an MCU control pin (CTRL1) to the output (P1) that is directed on the analog control of the HV power supply. The lines labeled TEST1A and TEST1B are connected to analog inputs of the MCU.

must be sacrificed to the USB module to be used as the $D\pm$ pins. Another two are employed by the embedded programmer port; these two could actually be shared with another application but since, as we will see, we won't need to stretch the pin availability we chose to keep the functions separate. Six pins are eaten up by the MCP voltage control therefore nine pins remain that can be safely dedicated to the multiscaler.

We opted to use eight of them to make a low frequency multiscaler in which each channel is counted by a one-bit counter (which is to say a flip-flop). The decision has been informed by our prospect of future requirements in the next upgrades of the apparatus, for which we expect the need of a high number of channels to be more likely than the need of high performance ones.

We chose to employ four 74HC74 chips as the required flip flops. These ICs implement the function of old glorious 5474/7474 (a pair of edge-triggered D-type flip flops), heritage of the 1970s, by employing the slightly newer CMOS technology. This means that unlike the 7474 they can be powered at 3.3 V and will provide CMOS-compatible output voltage levels that can be wired directly onto the MCU.

Before being fed into the 74HC74 the input signals need to be clamped to the IC supply rails. The IC provides by itself clamping diodes that grant that whenever as unbounded signal can be clipped to the supply rails by draining less than 20 mA the IC will do so. Although this feature paired with an input resistor might seem capable to provide the necessary clamping employing it will come at a cost: the excess positive voltage of an incoming signal will be shorted by the clamping diode onto the positive power supply rail and will cause a current that the power supply must be able to absorb. The 3.3 V is obtained from the 5 V USB by dropping the voltage through a specimen of TS2940, a factory-calibrated fixed voltage regulator that is not capable of absorbing negative currents re-enetering its output lead (moreover: it requires a minimum fixed load to provide proper regulation). The issues caused by a negative current entering the 3.3 V supply rail can be tempered by adding capacitance and zener diodes between the rail and the ground plane; experience dictates that regardless of these countermeasure discharging a positive signal onto the supply will cause rippling. This is especially bad for the MCP voltage regulation since the feedback lines are measured with reference to the 3.3 V rail.

To completely circumvent all of this we employed zener diodes referenced to the ground plane to clamp the signal. Since the ground plane is connected to the incoming coaxial cable shielding the clamping current is absorbed through the coaxial cable shielding and does not disturb the board supply voltages. A single input module for the multiscaler is shown in figure B.3.

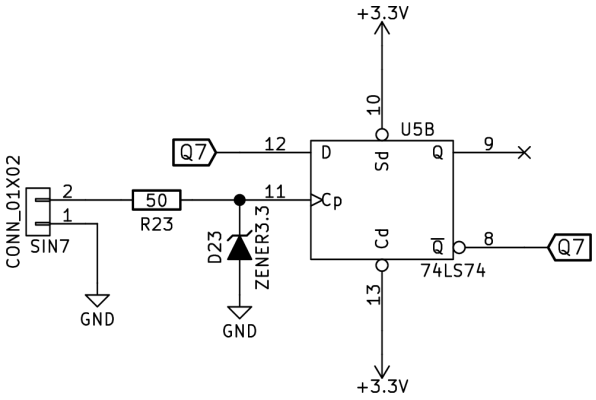


Figure B.3: The schematic of a multiscaler flip-flop input from the 50 Ω input line (SIN7) to the MCU input pin used to count the incoming pulses (Q7). The coupling to the coaxial cable is matched only for signals grater than the zener diode's threshold; small signals are partially reflected back onto the line. We verified that this is not an issue in our application.

B.3 Firmware

The Multiscaler firmware revolves around the multiscaler readout cycle which is the most time-critical function performed by the MCU. The multiscaler readout consists, in our implementation, of 79 assembler instructions. The selected MCU operates at a peak rate of 16 million instructions per seconds except when branching. Jump operation disrupt the preemptive op-code fetching pipe resulting in a double execution time. Our implementation contains 8 conditional jump instructions, each of which, when actually branching, skips the following 5 assembler instructions. Therefore the servicing of the multiscaler functions takes between 3.4 µs and 5.4 µs. If the MCU were to only service the multiscaler the maximum frequency of an evenly-spaced input signal that can be measured is 183 kHz.

Of course the firmware needs to attend other functions, namely driving the MCP outputs, measuring the output voltages and servicing the USB.

The MCP are driven by generating a 1 kHz square wave on the driving pins and skipping pulses whenever the output voltage after the first filtering stage is grater than the required setting. The servicing of the MCP driving runs once every millisecond, on even cycles it sets the outputs to zero, on odd cycles it sets them to one if and only if the last measured output voltage is below the requested threshold.

The USB host-to-device connection is serviced on interrupt. Whenever a frame is received from the host the USB hardware module raises an interrupt and a routine is invoked that processes the incoming frame. The device-to-host

communication is managed by making a USB frame always available in the output buffer[1]. Periodically the USB module's output buffer is checked and, if found empty, filled again with a frame containing the current device status. Consequence of this is that the data polled from the device will be always displaying the device status at the time the last frame was read out. This is not a problem in our application since we poll the status every second and suffering a one second latency time in the readout is not harmful to any of our operations, still in our driving software we included an option to perform a current status readout by simply performing two consecutive readouts.

During all of the aforementioned operations we interleaved the other operations with the multiscaler servicing so that no matter what operation is performed no 10 µs pass without a multiscaler readout. Taking into account the probability of collisions in incoming signals we expect the accuracy on the MCP rate channels, whose rate oscillates around 30 Bq to be $0.4 \cdot 10^{-3}$ and the accuracy of the beam intensity monitor which operates at about 900 Bq to be 1%. In neither of the two cases the accuracy of the MCU quartz, that provides the time reference, is a limiting factor. The reduction in the count rate induced by colliding signals arriving within the multiscaler servicing window is actually a systematic that can be compensated for, if needed, but this is not required in our case.

[1]Be aware that what we are referring to is the buffer that holds the data to be transmitted from the device to the host. The actual convention is to call that the input buffer since the USB protocol is generally described with the perspective of the host; still since here we are describing the operations from the perspective of the device we found clearer to refer to it as the output buffer.

A screenshot of Captor*'s web interface displaying a panel that allows average, comparison and subtraction of spectra acquired in different experimental conditions.*

Captor

That which can be destroyed by positrons, should be

Pulse Monitor | LOG OUT

ACQUIRE | COMBINE | DETREND | SHOOT | ABOUT

Show Samples

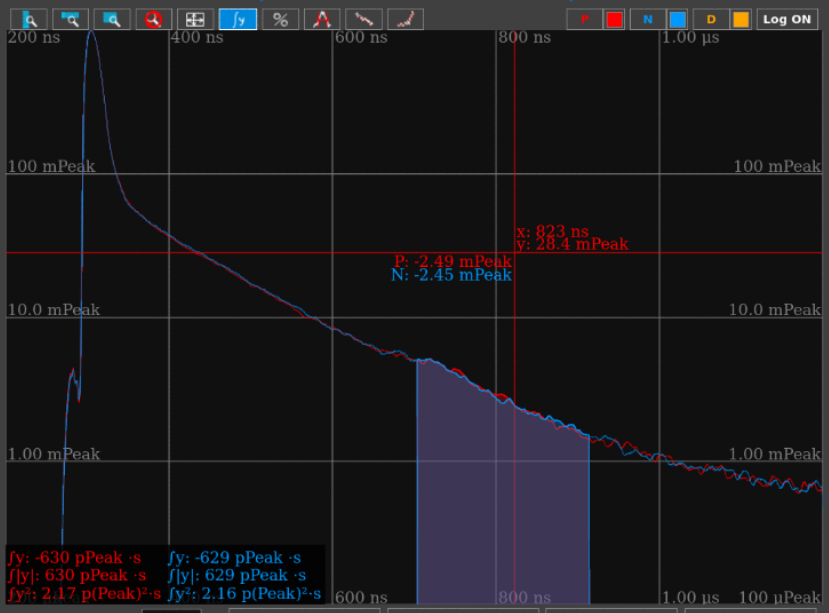
Create New Composition

- n=3/
- metastable/
  - metastabletest
  - metastable3polv
  - metastable2polhbis
  - metastable2polh

Header ver.: 5 File ver.: 6
Timestamp: 12/3/2016 11h 50m 27s 785ms
UniqueID: 14126750120959520 99
Hash: 95B7F80000020000
Metadata:

- metastable1
- dome/
- afterRug/
- UV/
- Transmission3cm/
- Transmission/
- TOF/
- Rydberg_2015/
- Rydberg/
- PsUVIR/
- IRscanpower/
- 2017/
- 2016/
- 2015/
- tst
- triggertest
- trans-test1
- testtriggerweplus
- test_switch_minikicker3
- test_switch_minikicker2
- test_switch_minikicker
- test_delay_kicker
- test
- synchronization_kicker
- statistic4.3PMT16vs20b
- statistic4.3PMT16vs20
- statistic3.3vs4.3lasersync
- statistic3.3vs1lasersync
- statistic2.1vs1lasersync
- signalPbF2
- positrononMCP
- newsource2
- delay195Vin1T
- cosmic3
- cosmic2
- cosmic
- checkfrequency4
- checkfrequency3
- checkfrequency2
- checkfrequency1
- checkfrequency
- UV73IR15b
- UV73IR15a
- UV+IRsynchro.final
- UV+IRsynchro.
- TriggerSetup19_07_2017
- Test_15_06_2017
- Test7_15_06_2017
- Test6_15_06_2017
- Test5_15_06_2017
- Test4_15_06_2017
- Test3_15_06_2017
- Test2_19_06_2017b
- Test2_19_06_2017
- Test2_15_06_2017
- Test-PosTarget-KickerOff
- T_100degrees3.3keVvs1keVb
- T_100degrees3.3keVvs1keV
- SyncTest091017
- RWok
- PulsesSynchroTest
- Ps_kicker3.3_10Kvs400K
- NewSource1
- NewComposition_15_08
- LaserSync090817
- Kickersync
- Kicker-3.3-keV
- HbarProd_12-09-2017
- HbarProd_11-09-2017
- HbarProd_1
- 43_PsRydberg_IR025CCW_PMT20_Recons
- 43_PsRydberg_IR025CCW_PMT20_FromBa
- 43_PsRydberg_IR025CCW_PMT20
- 29_RydbergExcitation_1708nm
- 28_PhotoionizationAtNight
- 27_Photoionization_newMod_PMT20
- 27_Photoionization_newMod_PMT16
- 27_Photoionization_PMT20
- 27_26_25_Photoionization_PMT20
- 27_26_25_Photoionization_PMT16
- 27_26_25_27Photoionization_PMT20
- 27_26_25_27Photoionization_PMT16
- 26_Photoionization_PMT20
- 26_Photoionization_PMT16
- 25_Photoionization_PMT20
- 25_Photoionization_PMT16
- 24_Photoionization_PMT20
- 23_Photoionization_Ground_PMT20
- 22_Photoionization_MiniKickerOn
- 21_Photoionization_MiniKickerOn_ResetCc
- 20_Photoionization_MiniKickerOn
- 20170811

Compositions:metastable/metastable2polh



Smearing (ns) 2 | Smear OFF | Auto Normalize ON | Intepolate OFF | Align OFF

Plot Combination

Positive

#0 2016_04_12/11_54_54_Lecroy_CH1234.bin — ON X
HiRes CH: NONE CH 1 CH 2 CH 3 CH 4
LoRes CH: NONE CH 1 CH 2 CH 3 CH 4
To normalize divide by: 1

#1 2016_04_12/11_54_54_Lecroy_CH1234.bin — ON X
HiRes CH: NONE CH 1 CH 2 CH 3 CH 4
LoRes CH: NONE CH 1 CH 2 CH 3 CH 4
To normalize divide by: 1

#2 2016_04_12/12_00_30_Lecroy_CH1234.bin — ON X
HiRes CH: NONE CH 1 CH 2 CH 3 CH 4
LoRes CH: NONE CH 1 CH 2 CH 3 CH 4
To normalize divide by: 1

#3 2016_04_12/12_00_30_Lecroy_CH1234.bin — ON X
HiRes CH: NONE CH 1 CH 2 CH 3 CH 4
LoRes CH: NONE CH 1 CH 2 CH 3 CH 4
To normalize divide by: 1

#4 2016_04_12/12_06_06_Lecroy_CH1234.bin — ON X
HiRes CH: NONE CH 1 CH 2 CH 3 CH 4
LoRes CH: NONE CH 1 CH 2 CH 3 CH 4
To normalize divide by: 1

#5 2016_04_12/12_06_06_Lecroy_CH1234.bin — ON X
HiRes CH: NONE CH 1 CH 2 CH 3 CH 4
LoRes CH: NONE CH 1 CH 2 CH 3 CH 4
To normalize divide by: 1

#6 2016_04_12/12_11_42_Lecroy_CH1234.bin — ON X
HiRes CH: NONE CH 1 CH 2 CH 3 CH 4
LoRes CH: NONE CH 1 CH 2 CH 3 CH 4
To normalize divide by: 1

#7 2016_04_12/12_11_42_Lecroy_CH1234.bin — ON X
HiRes CH: NONE CH 1 CH 2 CH 3 CH 4
LoRes CH: NONE CH 1 CH 2 CH 3 CH 4
To normalize divide by: 1

#8 2016_04_12/12_17_18_Lecroy_CH1234.bin — ON X
HiRes CH: NONE CH 1 CH 2 CH 3 CH 4
LoRes CH: NONE CH 1 CH 2 CH 3 CH 4
To normalize divide by: 1

#9 2016_04_12/12_17_18_Lecroy_CH1234.bin — ON X
HiRes CH: NONE CH 1 CH 2 CH 3 CH 4
LoRes CH: NONE CH 1 CH 2 CH 3 CH 4
To normalize divide by: 1

#10 2016_04_12/12_23_15_Lecroy_CH1234.bin — ON X
HiRes CH: NONE CH 1 CH 2 CH 3 CH 4
LoRes CH: NONE CH 1 CH 2 CH 3 CH 4
To normalize divide by: 1

#11 2016_04_12/12_23_15_Lecroy_CH1234.bin — ON X
HiRes CH: NONE CH 1 CH 2 CH 3 CH 4
LoRes CH: NONE CH 1 CH 2 CH 3 CH 4
To normalize divide by: 1

#12 2016_04_12/12_28_51_Lecroy_CH1234.bin — ON X
HiRes CH: NONE CH 1 CH 2 CH 3 CH 4
LoRes CH: NONE CH 1 CH 2 CH 3 CH 4
To normalize divide by: 1

#13 2016_04_12/12_28_51_Lecroy_CH1234.bin — ON X
HiRes CH: NONE CH 1 CH 2 CH 3 CH 4
LoRes CH: NONE CH 1 CH 2 CH 3 CH 4
To normalize divide by: 1

#14 2016_04_12/12_34_28_Lecroy_CH1234.bin — ON X
HiRes CH: NONE CH 1 CH 2 CH 3 CH 4
LoRes CH: NONE CH 1 CH 2 CH 3 CH 4
To normalize divide by: 1

#15 2016_04_12/12_34_28_Lecroy_CH1234.bin — ON X
HiRes CH: NONE CH 1 CH 2 CH 3 CH 4
LoRes CH: NONE CH 1 CH 2 CH 3 CH 4
To normalize divide by: 1

#16 2016_04_12/12_40_04_Lecroy_CH1234.bin — ON X
HiRes CH: NONE CH 1 CH 2 CH 3 CH 4
LoRes CH: NONE CH 1 CH 2 CH 3 CH 4
To normalize divide by: 1

Negative

#0 2016_04_12/11_52_06_Lecroy_CH1234.bin — ON X
HiRes CH: NONE CH 1 CH 2 CH 3 CH 4
LoRes CH: NONE CH 1 CH 2 CH 3 CH 4
To normalize divide by: 1

#1 2016_04_12/11_52_06_Lecroy_CH1234.bin — ON X
HiRes CH: NONE CH 1 CH 2 CH 3 CH 4
LoRes CH: NONE CH 1 CH 2 CH 3 CH 4
To normalize divide by: 1

#2 2016_04_12/11_57_42_Lecroy_CH1234.bin — ON X
HiRes CH: NONE CH 1 CH 2 CH 3 CH 4
LoRes CH: NONE CH 1 CH 2 CH 3 CH 4
To normalize divide by: 1

#3 2016_04_12/11_57_42_Lecroy_CH1234.bin — ON X
HiRes CH: NONE CH 1 CH 2 CH 3 CH 4
LoRes CH: NONE CH 1 CH 2 CH 3 CH 4
To normalize divide by: 1

#4 2016_04_12/12_03_18_Lecroy_CH1234.bin — ON X
HiRes CH: NONE CH 1 CH 2 CH 3 CH 4
LoRes CH: NONE CH 1 CH 2 CH 3 CH 4
To normalize divide by: 1

#5 2016_04_12/12_03_18_Lecroy_CH1234.bin — ON X
HiRes CH: NONE CH 1 CH 2 CH 3 CH 4
LoRes CH: NONE CH 1 CH 2 CH 3 CH 4
To normalize divide by: 1

#6 2016_04_12/12_08_54_Lecroy_CH1234.bin — ON X
HiRes CH: NONE CH 1 CH 2 CH 3 CH 4
LoRes CH: NONE CH 1 CH 2 CH 3 CH 4
To normalize divide by: 1

#7 2016_04_12/12_08_54_Lecroy_CH1234.bin — ON X
HiRes CH: NONE CH 1 CH 2 CH 3 CH 4
LoRes CH: NONE CH 1 CH 2 CH 3 CH 4
To normalize divide by: 1

#8 2016_04_12/12_14_30_Lecroy_CH1234.bin — ON X
HiRes CH: NONE CH 1 CH 2 CH 3 CH 4
LoRes CH: NONE CH 1 CH 2 CH 3 CH 4
To normalize divide by: 1

#9 2016_04_12/12_14_30_Lecroy_CH1234.bin — ON X
HiRes CH: NONE CH 1 CH 2 CH 3 CH 4
LoRes CH: NONE CH 1 CH 2 CH 3 CH 4
To normalize divide by: 1

#10 2016_04_12/12_20_27_Lecroy_CH1234.bin — ON X
HiRes CH: NONE CH 1 CH 2 CH 3 CH 4
LoRes CH: NONE CH 1 CH 2 CH 3 CH 4
To normalize divide by: 1

#11 2016_04_12/12_20_27_Lecroy_CH1234.bin — ON X
HiRes CH: NONE CH 1 CH 2 CH 3 CH 4
LoRes CH: NONE CH 1 CH 2 CH 3 CH 4
To normalize divide by: 1

#12 2016_04_12/12_26_03_Lecroy_CH1234.bin — ON X
HiRes CH: NONE CH 1 CH 2 CH 3 CH 4
LoRes CH: NONE CH 1 CH 2 CH 3 CH 4
To normalize divide by: 1

#13 2016_04_12/12_26_03_Lecroy_CH1234.bin — ON X
HiRes CH: NONE CH 1 CH 2 CH 3 CH 4
LoRes CH: NONE CH 1 CH 2 CH 3 CH 4
To normalize divide by: 1

#14 2016_04_12/12_31_39_Lecroy_CH1234.bin — ON X
HiRes CH: NONE CH 1 CH 2 CH 3 CH 4
LoRes CH: NONE CH 1 CH 2 CH 3 CH 4
To normalize divide by: 1

#15 2016_04_12/12_31_39_Lecroy_CH1234.bin — ON X
HiRes CH: NONE CH 1 CH 2 CH 3 CH 4
LoRes CH: NONE CH 1 CH 2 CH 3 CH 4
To normalize divide by: 1

#16 2016_04_12/12_37_15_Lecroy_CH1234.bin — ON X
HiRes CH: NONE CH 1 CH 2 CH 3 CH 4
LoRes CH: NONE CH 1 CH 2 CH 3 CH 4
To normalize divide by: 1

Appendix C

Captor

Nothing is real until you put it in the VCR.

– J.G. Ballard

Anybody who worked in the AE$\bar{g}$IS collaboration between 2015 and 2018 has at least once employed data acquired through *Captor* in her work, many have connected to *Captor* itself to view and analyze acquired data, some have connected to it to directly control the data acquisition process. To most users *Captor* appears as nothing more than a web interface that allows management, visualization and analysis of data acquired from the *Breadbox* detectors[1]. In this appendix we will take a look at what exactly is *Captor* and what happend behind the user interface.

Captor is a Debian-based acquisition server, that is a computer whose main task is to collect and store data from detectors. On top of this online data analysis capabilities have been added to the machine. All of the operations performed by *Captor* are run through a suite of sixteen executables whose operation can be performed entirely through the command line interface (CLI). These executables have been written entirely in C++ and are compiled and linked in a single-pass by direct inclusion of all of their source code. This technique has the main advantage of faciltating the GCC's optimizer job by allowing it to inline functions and sort routines to minimize jump distances; and the main disadvantage to increase the time and memory required by the compiling process. In our case since the source code of the *Captor* suite ranges around 17 thousand lines the compiling time equates approximately the

[1] And, in late 2017, some of the production trap detectors

linking time so no game-changing improvement in the compiling time is given by compiling separate object files and linking them afterwards; we therefore opted for a less efficient compiling in favor of more efficient executables.

The *Captor* command suite source code is based on an ecosystem of headers that provide some general capabilities that are required for the implementation of the CLI commands. More complex high-level capabilities are implented in headers that make use of the low-level headers. The source code of the various suite commands intersects since each command includes only the headers required for its implementation. If we were to draw the inclusion structure of the source code it would look like a tree whose roots are the CLI commands, which then fuse into a trunk consisting of the high-level functions, which then divides into the branches representing, in this metaphor, the low-level capabilities. The careful centralization of all of the functions whose coherence across the entire suite of commands is paramount to the correct operation of the suite itself (such as, to give an example, the magic numbers employed in the file headers) grants that punctual updating will always transversally affect the entire codebase and coherent behavior across the suite will be granted.

All of the *Captor* commands are equipped with a complete user guide that can be invoked through the CLI itself by operating the `-h` switch. The CLI guide is written within the C++ source code inside of properly marked comments, as is common with many documentation utilities e.g.: doxygen. A specifically written utility gathers all of these comments and compiles them into a C++ routine, different for each executable, that provides the help printing capability. This allows the user guide to be written side to side to the code that implements the interface to specific functionalities and allows much easier maintaining of an up-to-date documentation to the commands themselves.

C.1 File format

Data acquired through the *Captor* server is saved in a binary file format tailored to suit the necessities of the server. Different types of files are stored inside of the *Captor* database with each file type representing a different type of record we might want to store (e.g.: an oscilloscope curve, an image from the camera, etc.). Each file type is stored inside what internally is called a *Databank* that is a folder tree structure situated in the *Captor* database that stores all and only the files of that specific type. This allows us to define an URL structure for *Captor* files in the form:

`<databank>:<filepath>`

and employ it across the entire CLI suite and by natural extension, as we will see later on, within the web API interface.

All of the *Captor* files start with a header with a fixed structure. It begins

with an 8-byte-long magic number[2] followed by a 2-byte integer that contains the header format revision number. The header format revision number defines how the rest of the header should be interpreted; currently we are employing the fifth revision of the file header which contains a timestamp for the file creation, a 64-bit unique file identifier, a hash of the entire file content (header excluded), user-provided metadata (a string that the operators can add to the file to annotate the operations performed when acquiring the data) and a file format revision number. The file format revision number is then used to determine how the rest of the file content should be interpreted.

Every time a new feature was requested that required the addition of new fields to the *Captor* files, a new file format was defined and the revision number was incremented. The code that interprets the old format is kept inside of the codebase and a switch statement chooses the appropriate decoder based on the revision number found. Whenever files are saved it is required that they be saved always at the latest revision. This scheme presents several advantages: one is that it allows us to lazily update file formats (whenever we implement a new revision we are not required to update the entire *Databank* to the new format), another is that since files share the same header format it permits the implementation of *Databank*-agnostic versions of file management commands (such as search by hash or cloning with new unique id assignment) which allows us to greatly reduce the code redundance.

Due to the file format being subject to change and revision the user is never expected to interact directly with the base *Captor* file but, instead, to use the CLI command suite to interact with them.

C.2 Acquisition drivers

The current revision of *Captor* supports acquisition through an Hamamatsu frame grabber[3] and three oscilloscopes[4]. The oscilloscopes are read through their VISA-over-ethernet protocol, the frame grabber through the firewire; each acquisition type features a specialized executable to perform the operation.

Operation of the acquisition inside of *Captor* is performed by running an instance of the appropriate acquisition utility for each run of the apparatus. On startup the executable arms, if necessary, the oscilloscope or the camera, waits for data to be available on the device, acquires it, saves it (if requiested) into the appropriate *Databank* and exits. The suite features daemons that, at request, automate the acquisition process by starting new instances of the executables after each acquired shot.

[2]For the sake of pedantry one among 6071208897174521153, 5570745289071347009, 4848494792482964801 and 6071212195709404481 with the multiplicity of magic numbers being a vestigial feature.

[3]Hamamatsu Flash4 v

[4]Tektronix TDS 5054B, Lecroy HDO6104 and Lecroy waveRunner 104MXi-B

The oscilloscope acquisition drivers have been developed based on the libvisa opensource library and the three oscilloscopes' documentation. Unless requested the executable won't arm the oscilloscope for a new acquisition but instead download the curves already displayed on the oscilloscope screen; options are available to avoid saving the data if after download it is discovered to be not unique. The format in which data is saved preserves it in the original ADC scale and accompanies it with the offset and scale constants necessary to perform the conversion to the Volt vs second scale. In general *Captor* is designed to conserve the original data inside of its record and to decode it on the fly whenever required.

The frame grabber was supplied with its Windows-only proprietary DLL interface library and no indication about the underlying communication protocol. Usual practice in this circumstance is to create a Windows test application, dump the communications between the driver and the device, then deduce the protocol from that. Unfortunately intercepting the firewire port is not as easy as intercepting a serial or a USB port and moreover any attempt on our side to make the Windows driver work from inside a virtual machine encountered little success. In a turn of luck by analyzing the library dependencies[5], we were able to guess that the underlying protocol is not proprietary but indeed an implementation of the IIDC's standard DCAM protocol, with some of the features mapped in a slightly unusual way. We were, therefore, able to employ the libdc1394-22 opensource library to develop the proper drivers to operate the camera.

C.3 Web interface

The CLI interface constitutes a solid foundation upon which to build the functionalities of the server; at the same time most users would find annoying having to learn its usage and distracting having to employ it during data taking. We therefore implemented a web interface to smooth the learning curve of the tool and better the user experience. The main advantage of a web interface as opposed to a window-server-based graphical user interface (GUI) is that it allows user to remotely connect to *Captor* regardless of the operating system or architecture they are running, as long as they have an HTML5-compatible web browser.

The first ingredient that we need to develop the web interface is a set of web APIs to allow the execution of the backend CLI executables by invoking the appropriate URL. We based the web interface of *Captor* on Apache2 and its API interface on PHP. The *Captor* web API interface is basically a set of PHP wrappers. At this stage the main issue we face is security: we forced the APIs to be accessible only through the HTTPS protocol, made them

[5]Dependencies computed in "Library Explorer"

require the POST of the appropriate tokens to be run and severely limited the exposed capabilities in order to reduce the exploitation risk.

After the APIs have been exposed an HTML-JavaScript web interface has been developed. The *Captor* web interface offers the browsing of acquired images and oscilloscope curves, the computation of FWHM values and diameters of the beam spot. Images and oscilloscope curves can be combined through a variety of mathematical operations, RMS noise can be computed online and the integral of pictures or curves can be computed online too. To evaluate the presence of positronium in the SSPALS spectra the oscilloscope curve viewer offers logarithmic plot options and the capability of fitting curves with straight lines or offset exponentials (curves in the form $y = ae^{-b \cdot x} + c$). Any value (fits, integrals, interpolations, etc.) obtained through *Captor*'s interface is guaranteed to be publication-grade.

Captor's curve elaboration interface offers the possibility to gather curves into *Collections* that can be averaged; average curves can then be subtracted to evaluate difference in signal. Curves in *Captor* collections can be obtained by merging high-gain and low-gain acquisitions, can be normalized to their peak and if necessary can be automatically re-aligned horizontally. Since *Captor* curves can be acquired by different oscilloscopes, with different sampling frequency settings and be horizontally shifted, *Captor*'s curve arithmetic does not require identical sampling of the curves it is operating on. Unless this feature is disabled by the user, missing samples are automatically generated by cubic interpolation from the available points. At any step of the elaboration pipeline curves and images can be downloaded in a variety of formats (seven formats are supported for curves, four for images) for further processing using other software.

As of 2017 the *detrending* algorithm described in section 3.5.1 has been added to *Captor's* web interface and it is now common practice for the AEgIS operators to always employ detrending to evaluate signal differences.

Another key design choice when implementing *Captor*'s web interface was to endow every element of the interface with a mouse-over caption description. This allows operators with limited training to autonomously discover new features as they employ the tool, therefore reducing the burden on trained operators to babysit newcomers.

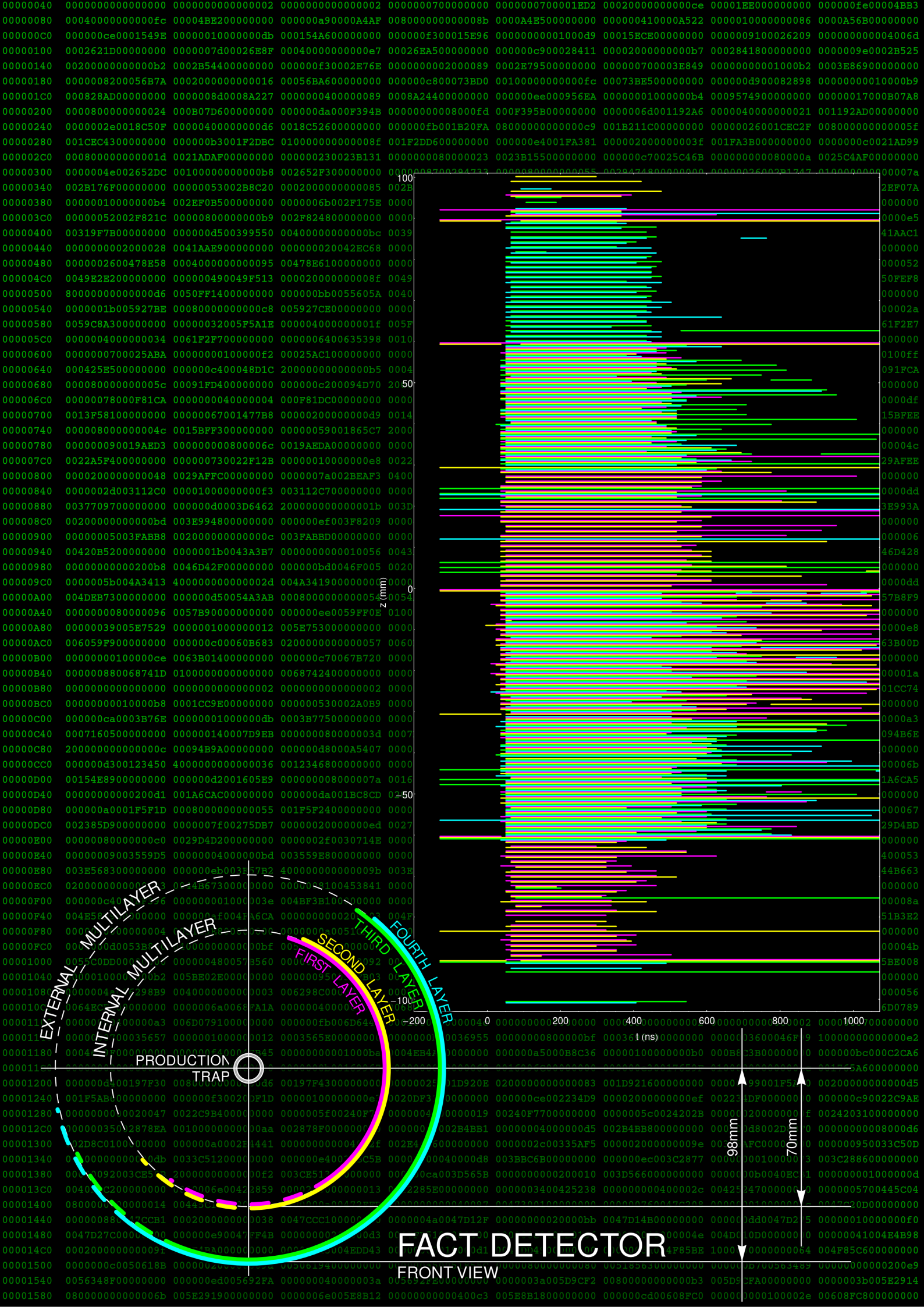
C.4 Prospects

When we started the implementation of *Captor* we did not envision the scope to which this tool would have evolved in time. Originally *Captor* was designed just as a data storage for spectroscopy experiments in the *Breadbox*. In time it has become a diagnostic and data analysis tool. Designed to be employed exclusively in the *Breadbox* it found applications in the commissioning of

electronics, in the characterization of HV pulsers and in the measurement of trigger delays and jitters.

Currently *Captor* is being used to align the positron beam onto the target in the production zone (the area surrounding the production trap inside of the 1 T magnet). The *Captor* server is an ongoing effort to ease operations in the AE$\bar{g}$IS apparatus with new features being added to the server to aswer to the need of the operators and the operators finding new uses for the server in response to the new features availability. A major upgrade of the server is scheduled within 2018 but all indications point toward the fact that it won't be the last.

In the background undecoded ethernet frames sent by one of the FPGAs of the FACT detector to the control computer. In the foreground a recording of a positron shot shot onto the production target located within the $1\,\mathrm{T}$ *magnet. Each line in the plot represent an activated fiber with the endpoints representing the timing of the rising and falling edge of the fiber. Each y coordinate corresponds to a different fiber; fibers have been sorted so that the approximate z position of each fiber can be read in the y scale of the plot. The curvature on the left side of the main signal bulk shows the different time at which different fibers are hit by* γ *rays due to time of flight. The different time-over-threshold is indicative of the different* γ *ray flux different fibers are receiving due to the solid angle they cover with respect to the target.*


FACT DETECTOR
FRONT VIEW
EXTERNAL MULTILAYER
INTERNAL MULTILAYER
FIRST LAYER
SECOND LAYER
THIRD LAYER
FOURTH LAYER
PRODUCTION TRAP
z (mm)
t (ns)
100
50
0
−50
−100
−200
0
200
400
600
800
1000
98mm
70mm

Appendix D

Drivers of FACT

I don't want to rule the universe. I just think it could be more sensibly organized.

– Eliezer Yudkowsky
Harry Potter and the Methods of Rationality

The FACT detector was commissioned in 2014[124] by the original FACT team which then moved on onto different projects. When commissioning the detector the original work group developed a ROOT-based control software that was delivered alongside the detector. The original software required an X11 window server to be run, was able to connect to the detector only through serial port connection, required an operator to always re-arm the detector before each run and could only save data locally. It was also riddled by a series of severe bugs, the worst of which was that, were the operator to perform the wrong procedure to close the program, all of the acquired data from the current session would be lost. Although the original drivers made an excellent work in allowing the first tests on the detector hardware, they would poorly serve the goal of allowing the precise characterization of the detector response and its employment during data taking.

To make the use of FACT possible during the $\bar{\mathrm{H}}$ runs of late 2017 we had to develop a driver program suited for the actual data taking.

D.1 Design choices

In the winter of 2016 we started developing a new driver program for the FACT detector with the design goals for it to be efficient, thoroughly debugged, executable without a graphic server and autonomous (that is capable of being

made to run without an operator). Finally the new drivers would need to be capable of automatically calibrating the threshold of each fiber to match a target dark count rate as well as to automatically send FACT data to the centralized AE$\bar{g}$IS DAQ server.

We started by extracting the code implementing the interface classes from the original driver and adapted it to be compiled within a standalone executable (as opposed as to be part of ROOT-compiled libraries). We completely removed every dependency of these classes from the ROOT libraries and proceeded to hardcode most of the support files in the executable source code. Since with no doubt this last sentence will have many programmers shrivel let's expand a little on the reasons behind this choice.

Whenever a program needs in its operation knowledge of settings such as hardware wiring, input and output paths or calibration constants it is a widely accepted good practice to have the program load them from a settings file. The main benefits of this design are the production of a lighter executable (if the required data is in considerable amount) and to allow the user to change these settings without having to compile again the executable. The main disadvantage of this pattern is that the correct working of the executable becomes dependent on the existence and correct formatting of several settings files. Shall any of these files be damaged or moved, the executable will stop working and, were the number, position and exact format of the settings files not fully documented, the user might be put in the position of not being able to operate the program at all. The main considerations to be done when choosing one of the two patterns are: how likely are the settings to be changed (it is unlikely for the value of π to need to be placed in an editable settings file) and what is the expertise level required of a user to properly determine correctly new settings after a change is needed (if only overly expert users can properly edit the settings then they are also capable of recompiling the code, moreover if editing settings presents risks then it is better to limit the possibility to do so for non expert users). In the case of the FACT drivers we settled on having two external editable setting files, on hardcoding every setting whose change required the FPGA firmware to be edited accordingly and exposing letting any user-available option through CLI parameters.

Another design choice was to simplify the operation of the detector to the point in which invoking the executable in the CLI without providing any parameter would arm the detector and, upon trigger arrival, result in the data being acquired, saved onto disk and sent to the centralized DAQ. Default values for all available options are hardcoded inside of the FACT driver executable and can be overridden by an external settings file (whose existence is not required for the execution of the driver). Moreover the CLI interface and the settings and output file formats are fully documented with the documentation hardcoded into the executable. Therefore a user having access to the sole compiled executable can employ it to set up completely the

environment required to successfully run it.

D.2 The serial port era

By design the FACT detector should have been operated through 17 FPGAs, each equipped with a 1Gbit Ethernet connection and an onboard buffer capable of storing 8 Mi ($= 8 \cdot (1024)^2$) 88-bit words. Each word is capable of storing the status of the 48 fibers managed by the FPGA (6 bytes), an additional byte used for debug purposes, a 32-bit timestamp and a one-byte-long checksum. Unfortunately the original FACT team delivered the detector with a firmware capable of employing only the serial port to communicate and 4 Ki ($= 4 \cdot 1024$) of the 8 Mi available words. This greatly slows down the operation of the detector and severely limits its maximum acquisition time: employing the original firmware a new word was produced for every tick of the onboard FPGA clock, which runs at 200 MHz, setting the maximum acquisition length through FACT to less than 20.5 µs.

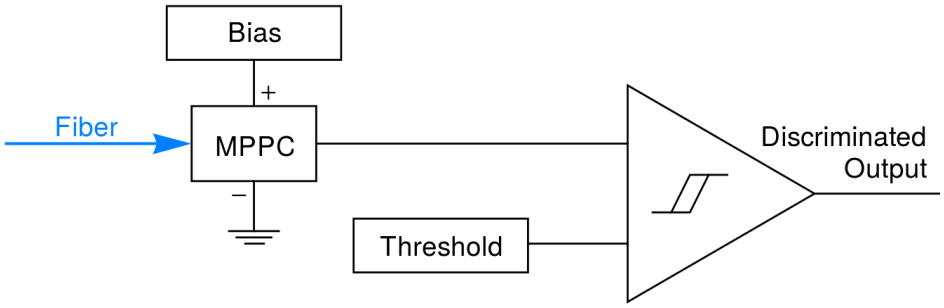


Figure D.1: Representation of the roles of the *Bias* and *Threshold* settings in the FACT discriminator. The *Bias* value determines a correction to the MPPC's supply voltage whereas the *Threshold* setting determines the value against which the MPPC output is compared to produce the discriminated output.

In the first configuration we ran the FACT detector in, the FPGA serial ports are connected to the control computer through 17 USB-to-Serial interfaces, which are connected to two 10-port USB 3.0 hubs which in turn are connected to USB 3.0 ports available on the control computer motherboard (see figure D.2). The original software used to communicate with a single FPGA at a time which turned into long operation times. We in turn took advantage of the faster USB connection and the bottleneck being located at the level of the individual FPGAs by operating simultaneously all of the FPGAs. The new FACT driver program operates in a multithread fashion with a thread dedicated to each operated FPGA and a coordinating thread

managing the interface with the user. This scheme offers the possibility of employing longer timeouts times and higher numbers of retries, since the malfunctioning of a single FPGA won't slow down the operation of the rest of the detector.

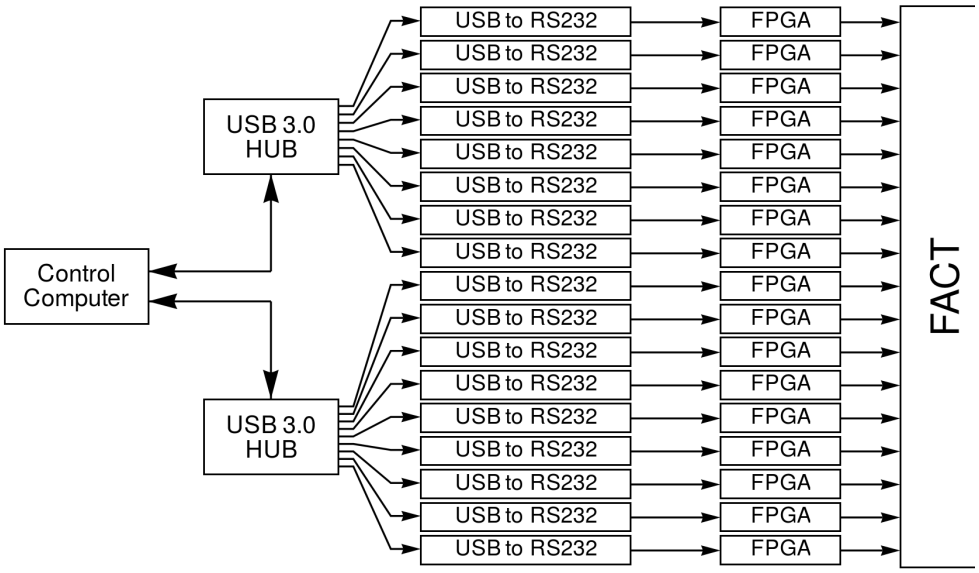


Figure D.2: Connection scheme of the FACT detector in the serial port configuration.

In parallel to the driver program we developed a calibration utility to characterize the detector's dark count rate as a function of the *Bias* and *Threshold* values. The output of each fiber is dependent on two FPGA-controlled parameters which can be set for each fiber. The *Bias* value, comprised between 0 and 127 determines a (small) change in the supply voltage of the fiber's MPPC. The *Threshold* value, comprised between 0 and 63, determines the value that the fast discriminator will compare the MPPC's output to to discriminate the fiber signal (see figure D.1 for a schematic representation of the process). The first iteration of the calibration utility scanned every combination of the *Bias* and *Threshold* values and recorded the activity of the fiber in a fixed time window. All of the detector fibers were scanned at the same time, nonetheless due to the extremely slow communication the characterization of the entire detector response took several hours.

These characterizations of the detector response evidenced errors in the detector wiring that caused the signal recorded by some fibers to be nonsense. See for example the plots in figure D.3 which in a correct working setting should have the shape of two distinct areas separated by a straight line. Correcting the wiring and installing additional shielding solved the problem.

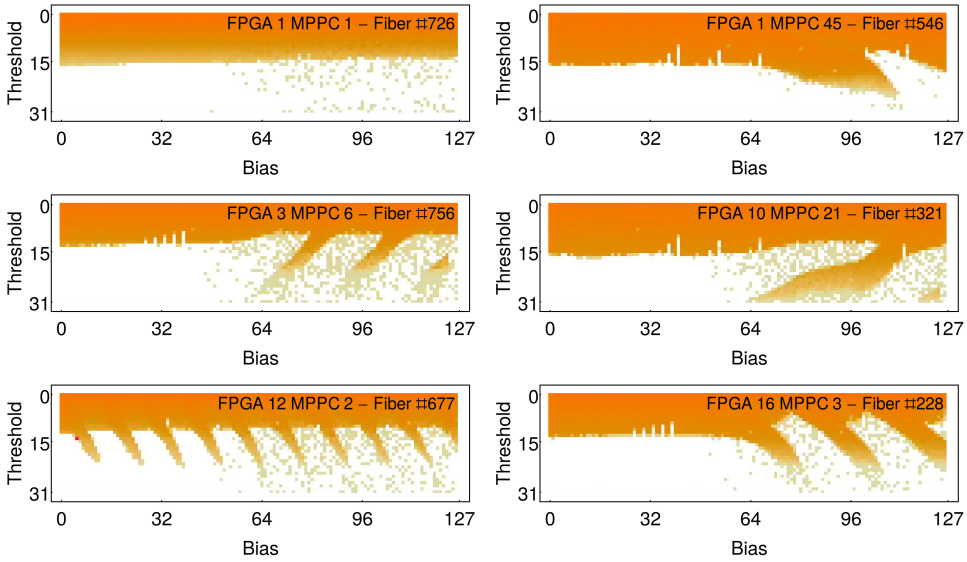


Figure D.3: Activity of the FACT fibers as a function of the *Bias* and *Threshold* values. On the top left the response we would expect within nominal operation conditions, in the other five plots the response caused by the incorrect wiring within the detector.

D.3 The ethernet era

In early 2017 we started upgrading the firmware of the FPGAs themselves. The first improvement we introduced was the addition of data compression in the form of saving new words only when we saw fibers changing status. The change increases the recording time of the detector by a factor that depends on the operations performed in AE$\bar{g}$IS at trigger time; we were able to observe it range from 10 to 10^6. The second improvement we introduced was enabling the ethernet interface to communicate with the FPGAs; in the perspective of trying to introduce as little changes as possible to the firmware most of the ethernet protocol was modeled around the serial protocol with the addition of a few new commands, the new connection configuration is shown in figure (D.4). Thirdly the onboard buffer was slightly increased passing from being 4096 to being 5632 words long.

On the driver side the multithread architecture has been maintained with the addition of a centralized packet receiver. The round trip for handshake packets leaving the host computer and bouncing back from the FPGAs has been measured to be around 12 µs and the typical handling time on the host side (which can be performed in parallel for packets coming from different FPGAs) is around 30 µs. Due to the response time of the FPGAs being predictably constant and the ethernet protocol not being arbitrated we initially

experienced frequent packet collisions. We solved the problem by staggering the starting time of the FPGA-handling threads by 20 µs and introducing randomized wait times after communication errors; these simple measures have made packet collisions on FACT practically nonexistent.

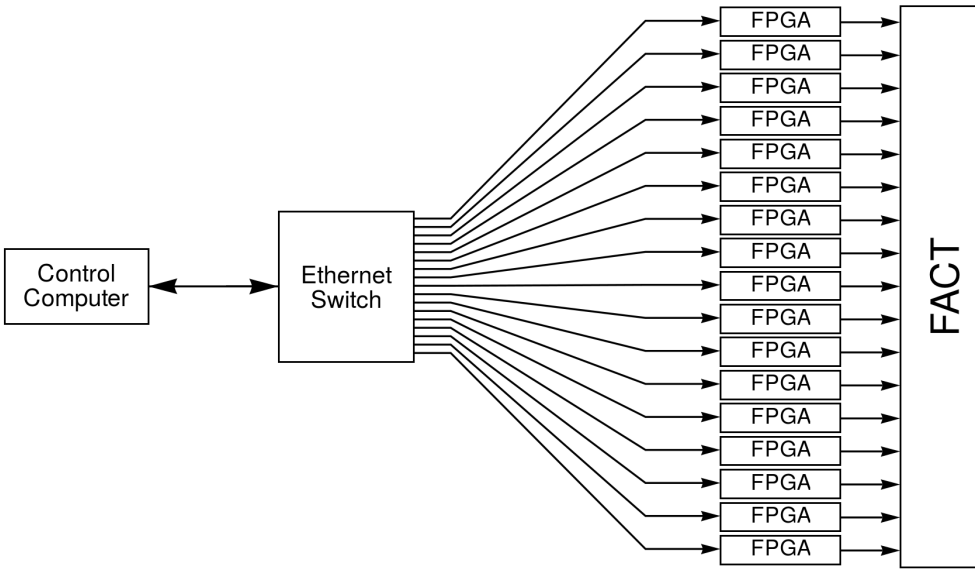


Figure D.4: Connection scheme of the FACT detector in the serial port configuration.

Two new acquisition modes had been developed. Instead of being able to only arm the detector and wait for data once, the operator can now ask the detector drivers to download data from the FGPAs' buffers as quickly as possible and to recycle immediately the data space, we call this operation mode the *Multiple* acquisition. Alternatively the driver can now instruct the FPGAs to spontaneously send data across the ethernet connection as soon as it is received; we call this acquisition mode *Stream*. Our early estimates say that, as long as the average rate of the generated words does not exceed 1 MBq per FPGA both the *Stream* and the *Multiple* should be able to record all of the produced data with the main difference among them being that the *Stream* mode should be able to run at higher rates but is less resistant to packet collisions than the *Multiple* mode.

A new algorithm was developed to find suitable *Bias* values to obtain the desired dark count rate. The algorithm worked by calibrating in parallel all of the FPGAs but each fiber in them sequentially. The algorithm employed the buffer filling as a measure of the activity on each fiber, therefore it was impossible for it to calibrate more than one fiber per FPGA at once; this set the running time of the calibration algorithm at 5 minutes. This algorithm

relied on bisection to search for the appropriate *Bias* setting, which is a technique that can land on a completely wrong setting if a single noisy data point is collected during the execution; this is not a capital issue but is bound to happen for a couple of fibers after each calibration.

Since the operation of the AEḡIS apparatus cycles at a rate of about one run every two minutes, a 5 minute long calibration procedure means that FACT must either be calibrated outside of the shift time or that operations in AEḡIS need to be stopped for a few cycles to allow for the detector calibration. At the same time calibrating frequently the detector is necessary due to the MPPC response being extremely sensitive to temperature with the dark count drifting by approximately a factor of two for each degree variation on the MPPC temperature.

To obtain a faster and more robust calibration procedure that could be performed without interrupting the operation of the rest of the experiment in fall 2017 we developed the algorithm described in the following paragraph.

D.4 Calibrating FACT

We want to find a *Bias* setting for the fibers of FACT that produces a certain rate of dark count (typically 100 Bq). Since FACT fibers' efficiency is not high and the detector redundancy limited, maintaining the highest possible *Bias* values is required if we aim to reconstruct vertexes; sub-optimal *Bias* values would likely degrade the signal to the point at which no reconstruction would be possible. Conversely excessive *Bias* values would on one side render the reconstruction ambiguous due to noisy fibers firing, on the other render the detector less efficient due to the blind time fibers are subject shortly after their activation.

Let's start by stating an observation method for the dark rate given by a certain *Bias* setting. We set the *Bias* level, then allow FACT to record noise for up to a certain interval of time, then observe how many events happened in the given time interval. Of course during the observation time the FPGA event buffer might saturate. Let n be the number of recorded events and Δt be the time elapsed between the beginning of the recording and the time at which the last event was recorded (see the scheme in figure D.5). If $n \neq 0$ we can write a tentative assessment on the fiber's dark rate as:

$$R = \frac{n \pm \sqrt{n}}{\Delta t} \tag{D.1}$$

Whereas if no event was recorded we can tentatively assert:

$$R \leq \frac{1}{\Delta t} \tag{D.2}$$

Multiple observations can be combined together, by adding the observed event count and the observation times, to better the precision of the rate assessment.

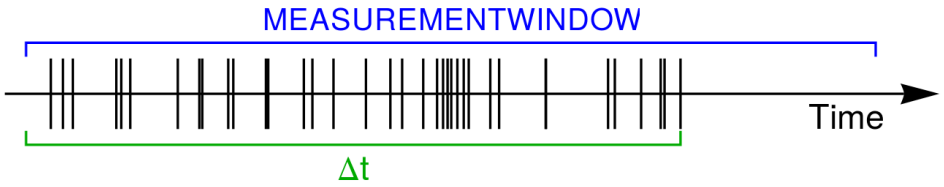


Figure D.5: After a recording has been performed the buffer might, due to saturation, display a trailing time devoid of events; therefore we employ the time up until the last event as the measurement time Δt instead of the full window width to avoid bias due to the buffer saturation.

Employing an arbitrary start time does not bias the measurement since the time between the arbitrary start and the first recorded event will be distributed like all of the other time intervals between successive events (the tail of the exponential distribution being an exponential distribution itself). Instead due to the possibility of buffer saturation the time separating the last event from the end of the recording time is not independent on the recorded data, therefore it is important to measure Δt only up until the last recorded event.

As long as we stick precisely to this procedure we are allowed to measure the dark count of multiple fibers at once since the effect of a noisy fiber filling up the buffer will be at most that of reducing the observation time for all of the others, but will not bias the assessed rates.

Our procedure to calibrate the fibers is now delineating, we just need to determine which *Bias* values we should be setting each fiber to before performing a rate observation; we will base our choice on the previous assessments on the R value for the fiber. At any point of the calibration process we will have an estimation of the noise rate of any given fiber for a certain number of *Bias* settings and no information about its behavior at other *Bias* values. First of all if we have no measurement at all (it is the first iteration of the algorithm) we proceed to measure the rate for the lowest available *Bias* value; if we have only one measurement of R we proceed to measure the rate for the highest available *Bias* setting. If we have two or more measurements we use them to generate an assessment of the probable dark count of the fiber for each possible *Bias* value. We do so by linearly interpolating the mean and uncertainty of the measuements preceding and following each measured *Bias* setting.

We then associate to each *Bias* setting an estimator of how "interesting"

that *Bias* setting is to us. Originally we meant this measure to be a measure of how likely we found that the specific *Bias* setting could be the one closest to giving the desired dark noise. This would have called for a Poissonian or Normal distribution being employed; instead we found that in simulated scenario the best performing "interest" estimator is a quadratic law:

$$I = \left(\frac{\Delta R}{R - T} \right)^2 \tag{D.3}$$

Where I is the interest estimator, T is the goal dark noise activity for the specific fiber, R is the rate assessment and ΔR its uncertainty. We determine which *Bias* setting presents the highest value of I and randomly pick between it, its preceding and its following values the *Bias* value that will be measured at the next iteration. This random factor is necessary since it may happen that, due to bad luck, the best *Bias* setting is observed to lay farther away from the target activity than other sub-optimal *Bias* values, in which case the optimal value would not be measured again. We show in figure D.6 an ongoing simulated calibration process in which both the interpolated assessment and the interest estimator I are displayed.

This algorithm has shown to perform well both in simulated contexts and when applied on the actual detector. When applied to FACT the algorithm typically converges to the best setting within 12 iterations; we settled to have the algorithm run 24 iterations for a typical detector calibration; with this setting the calibration execution time clocks at 4.6 s, which for frequent re-calibration of the fibers *Bias* settings.

The possibility to calibrate *Bias* values at every run of AE$\bar{\text{g}}$IS allows us to increase the reproducibility of the measurements and to safely allow higher levels of dark counts since we can ensure that any thermal drift that could increase them and degrade the data will not have time to take place before the next calibration procedure is performed.

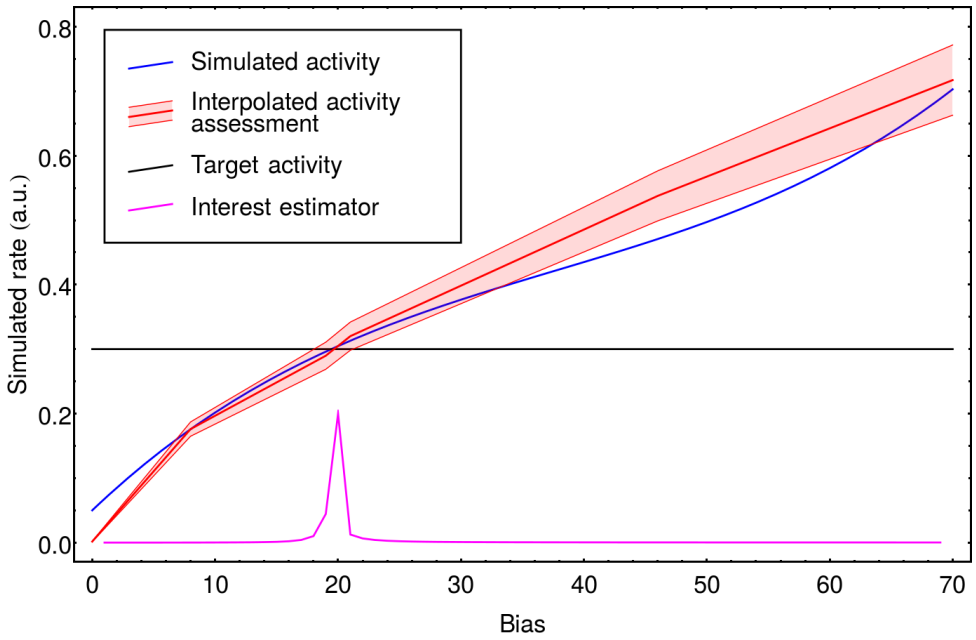


Figure D.6: An ongoing calibration process (here we display the sixth iteration) for a single fiber is being simulated. In blue we plot the simulated fiber activity which in a real case scenario tends to be much steeper. The red envelope is the interpolated assessment of R, the horizontal black line shows the target activity T. Finally in magenta we show (scaled down) the interest estimator I which has already converged to the optimal *Bias* setting

Bibliography

[1] Carl D. Anderson, *Physical Review* **43** 491 (1933)

[2] C.F. Marvin, *U.S. Department of Agriculture, Weather Bureau* "Climatological Data: California", Volume 35, (1933)

[3] W.H. Pickering, "Biographical Memoir of Carl David Anderson" National Academies Press, Washington D.C. (1998)

[4] C.D. Anderson and S.H. Neddermeyer *Physical Review* **50** 263 – 271 (1936)

[5] E.J. Williams and G.E. Roberts, *Nature* **145**, 102 – 103 (1940)

[6] S. De Benedetti, C.E. Cowan and W.R. Konneker, *Physical Review* **76** 440 (1949)

[7] R. Bjorklund, W.E. Crandall, B.J. Moyer and H.F. York, *Physical Review* **77** 213 – 218 (1950)

[8] M. Deutsch, *Physical Review* **82** 455 (1951)

[9] M. Deutsch and E. Dulit, *Physical Review* **76** 601 (1951)

[10] W.H. Sweet and G.L. Brownell, Localization of brain tumors with positron emitters *Nucleonics,* **11** 40 – 45 (1953)

[11] O. Chamberlain, E. Segerè, C. Wiegand and T. Ypsilantis, *Physical Review* **100** 947 (1955)

[12] C.S. Wu, *Pysical Review* **105** 1413 (1957)

[13] G. Lüders, *Annals of Physics* **2** 1 – 15 (1957)

[14] R.E. Bell, M.H. Jørgensen, *Canadian Journal of Physics* **38** 652 – 664 (1980)

[15] G.L. Brownell, *"A History of Positron Imaging"*, Presentation in Celebration of the 50th Year of Services of the Author to the Massachusetts General Hospital (1999)

[16] A.P. Mills Jr., P.M. Platzman and B.L.Brown, *Physical Review Letters* **41**, 1076 – 1079 (1978)

[17] P.B. Schwinberg, R.S. Van Dyck and H.G. Dehmelt, *Physics Letters A* **81** 119 – 120 (1981)

[18] R. Klapisch, *Hyperfine Interactions* **24 − 26** 311 – 319 (1985)

[19] S. Chu, A.P. Mills Jr. and John L. Hall, *Physical Review Letters* **52** 1689 – 1692 (1984)

[20] G. Bardin et al., *Physics Letters B* **137** 135 – 140 (1984)

[21] C.M. Surko, M. Leventhal and A.Passner Physical Review Letters **62** 901 – 904 (1989)

[22] R. Carosi et al., *Physics Letters B* **237** 303 – 312 (1990)

[23] G. Baur et al., *Physics letters B* **368** 251 – 258 (1996)

[24] M. Amoretti et al., *Letters to Nature* **419** 456 – 459 (2002)

[25] G.B. Andersen et al. (ALPHA Collaboration), *Nature Physics* **7** 558 – 564 (2011)

[26] M. Hori et al., *Nature* **475** 484 – 488 (2011)

[27] (ALPHA Collaboration) and A. E. Charman, *Nature Communications* **4** 1785 (2013)

[28] T.A. Wagner, S. Schlamminger, J.H. Gundlach, and E.G. Adelberger, *Classical and Quantum Gravity* **29** 184002 (2012)

[29] G. Drobychev et al, CERN-SPSC-2007-017 (2007), `http://cdsweb.cern.ch/record/1037532`

[30] S. Aghion et al. (AE$\bar{\text{g}}$IS collaboration), *Nature Communications* **5** 4538 (2014)

[31] A. Kellerbauer et al. (AE$\bar{\text{g}}$IS collaboration), *EPJ Web of Conferences* **126** 02016 (2016)

[32] R. Caravita et al. (AE$\bar{\text{g}}$IS collaboration), *Acta Physica Polonica B* **48** 1583 – 1592 (2017)

[33] D. Krasnický, R. Caravita, C. Canali, and G. Testera, *Physical Review A* **94** 022714 (2016)

[34] R.S. Brusa et al. (AE$\bar{\text{g}}$IS collaboration), *Journal of Physics: Conference Series* **791** 012014 (2017)

[35] F. Anderegg, E.M. Hollmann, and C.F. Driscoll, *Physical Review Letters* **81** 4875 – 4878 (1998)

[36] E. Vliegen and F. Merkt, *Journal of Physics B: Atomic Molecular and Optical Physics* **39** L241 (2006)

[37] E. Vliegen, S.D. Hogan, H. Schmutz, and F. Merkt, *Physical Review A* **76** 023405 (2007)

[38] A. Zecca, M. Bettonte, J. Paridaens, G.P. Karwasz, R.S. Brusa, *Measurement Science and Technology* **9** 409 – 416 (1998)

[39] L. Liszkay et al. *Journal of Physics: Conference Series* **443** 012006 (2013)

[40] C. Hugenschmidt, C. Piochacz, *Journal of large-scale research facilities* **1** A22 (2015)

[41] Radioactive Decay Data Tables, David C. Kocher Office of Scientific and Technical Information U.S. Department of Energy DOE/TIC-11026 **Apr** (1981)

[42] A.P. Mills Jr. and E.M. Gullikson, *Applied Physics Letters* **49** 1121 (1986)

[43] S.G. Karshenboim, *Physics Reports* **422** 1 (2005)

[44] A. Deller, A.M. Alonso, B.S. Cooper, S.D. Hogan, and D.B. Cassidy, *Physical Review A* **93** 062513 (2016)

[45] P.J. Schultz and K.G. Lynn, *Reviews of Modern Physics* **60** 701 (1988)

[46] A.P. Mills Jr. and W.S. Crane, Physical Review Letters **53** 2165 (1984)

[47] R.S. Brusa et al. (AEḡIS collaboration), *Acta Physica Polonica* **113** 1301f (2008)

[48] A.P. Mills Jr., Physical Review Letters **41** 1828 (1978)

[49] D.W. Gidley, W.E. Frieze, T.L. Dull, A.F. Yee, E.T. Ryan and H-M Ho, *Physical Review B* **60** R5157 (1999)

[50] M.P. Petkov, M.H. Weber, K.G. Lynn and K.P. Rodbell, *Applied Physics Letters* **79** 3884 (2001)

[51] C.L. Wang, M.H. Weber, K.G. Lynn and K.P. Rodbell, *Applied Physics Letters* **81** 4413 (2002)

[52] R.S. Yu, T. Ohdaira, R. Suzuki, K. Ito, K. Hirata, K. Sato, Y. Kobayashi, and J. Xu, *Applied Physics Letters* **83** 4966 (2003)

[53] L. Liszkay et al., *Applied Physics Letters* **92** 063114 (2008)

[54] S. Mariazzi, P. Bettotti, S. Larcheri, L. Toniutti, and R.S. Brusa, *Physical Review B* **81** 235418 (2010)

[55] S. Mariazzi, P. Bettotti, and R.S. Brusa, *Physical Review Letters* **104** 243401 (2010)

[56] R.S. Brusa, A. Dupasquier, Proceedings of the International School of Physics "Enrico Fermi" (2009) DOI 10.3254/978-1-60750-646-1-245

[57] S. Mariazzi, A. Salemi, and R.S. Brusa, *Physical Review B* **78**, 085428 (2008)

[58] J. Dryzek, P. Horodek, *Nuclear Instruments and Methods in Physics Research B* **266** 4000 – 4009 (2008)

[59] A.F. Makhov, *Soviet physics, Solid state* **2** 1934-41 (1960)

[60] A.F. Makhov, *Soviet physics, Solid state* **2** 1942-4 (1960)

[61] A.F. Makhov, *Soviet physics, Solid state* **2** 1945-51 (1960)

[62] E. Soininen et al., *Physical Review B* **46** 20 (1992)

[63] J. Algers et al., *Physical Review B* **67** 125404 (2003)

[64] S.Valkealahti and R.M. Nieminen, *Applied Physics A* **32** 95 – 106 (1983)

[65] S.Valkealahti and R.M. Nieminen, *Applied Physics A* **35** 51 – 59 (1984)

[66] J. Sempau et al., *Nuclear Instruments and Methods in Physics Research B* **207** 107 – 123 (2003)

[67] J. Baró et al., *Nuclear Instruments and Methods in Physics Research B* **100** 31 – 46 (1995)

[68] S. Agostinelli et al., *Nuclear Instruments and Methods in Physics Research A* **506** 250 – 303 (2003)

[69] J.M. Fernández-Varea, *Nuclear Instruments and Methods in Physics Research B* **73** 447 – 473 (1993)

[70] S.K. L. Sjue, F.G. Mariam, F.E. Merrill, C.L. Morris, and A. Saunders, *Review of Scientific Instruments* **87**, 015110 (2016)

[71] J.M. Fernández-Varea, D. Liljequist, S. Csillag, R. Räty, F.Salvat, *Nuclear Instruments and Methods in Physics Research B* **108** 35 – 50 (1996)

[72] V. J. Ghosh, G.C. Aers, *Physical Review B* **51** 45 – 59 (1995)

[73] J.A. Treurniet, D.W.O. Rogers, *NRC Report PIRS* **Oct** 669 (1999)

[74] R.M. Nieminen and J. Oliva, *Physical Review B* **22** 2226 (1980)

[75] R. Krause-Rehberg, H.S.Leipner, "Positron annihilation in Semiconductors – defect studies" *Springer* (1999)

[76] M.J.Puska, C.Corbel and R.M. Nieminen, *Physical Review B* **41** 9980 (1990)

[77] J. Dryzek and A. Czapla, *Journal of Physics: Condensed Matter* **10** L547 (1998)

[78] W. Egger, "Pulsed low-energy positron beams in material sciences" Physics with Many Positrons, Proceedings of the International School of Physics "Enrico Fermi" Course CLXXIV

[79] Y. Nagashima et al., *Physical Review B* **58** 012676 (1998)

[80] R. Paulin, G. Ambrosino, *Proceedings of the Conference held at Wayne State University* July 27 – 29 (1965)

[81] W.C. Sauder, Journal of Research ofthe Mational Bureau of Standards **72A** 1 (1968)

[82] M. Skalsey et al., *Physical Review Letters* **80** 3727 (1998)

[83] M. Skalsey et al., *Physical Review A* **67** 022054 (2003)

[84] L.D. Landau, E.M. Lifschitz, (1960) Course of Theoretical Physics - Volume I, Mechanics *Pergamon Press*

[85] Y. Nagashima et al., *Physical Review A* **52** 1 (1995)

[86] H. Saito and T. Hyodo, "New Directions in Antimatter Chemistry and Physics", Chapter 7, Kluwer Academic Publishers (2001)

[87] H. Saito and T. Hyodo, *Physical Review B* **60** 11070 (1999)

[88] R. Caravita, "Towards measuring gravity on neutral antimatter, PhD thesis", Ph.D. Thesis, Università degli studi di Genova (2017)

[89] S. Takada et al., *Radiation Physics and Chemistry* **58** 781 – 785 (2000)

[90] C. Dauwe et al., *Journal of Radioanalytical and Nuclear Chemistry* **210** 293 (1996)

[91] S. Mariazzi, L. Di Noto, G. Nebbia and R.S. Brusa, *Journal of Physics: Conference Series* **618** 012039 (2015)

[92] L. Di Noto, S. Mariazzi, M.Bettonte, G. Nebbia and R.S. Brusa, *The European Physical Journal D* **66** 118 (2012)

[93] S .Mariazzi et al., *Journal of Physics: Conference Series* **443** 012061 (2013)

[94] P. Crivelli, U. Gendotti, A. Rubbia, L. Liszkay, P. Perez and C. Corbel, *Physical Review A* **81** 052703 (2010)

[95] R.S. Brusa, G.P. Karwasz, M. Bettonte, A. Zecca, *Applied Surface Science* **116** 59 – 62 (1996)

[96] C. Macchi, S. Mariazzi, G.P. Karwasz, and R.S. Brusa, *Physical Review B* **74** 174120 (2006)

[97] C. Macchi et al.. *Physical Review B* **85** 214117 (2012)

[98] S. Aghion et al. (AEḡIS collaboration), *Nuclear Instruments and Methods in Physics Research B* **407** 55 – 66 (2017)

[99] E. Schreiber, H.J. Fitting, *Journal of Electron Spectroscopy and Related Phenomena* **124** 25 – 37 (2002)

[100] I.A. Glavatskikh, V.S. Kortov, and H.J. Fitting, *Journal of Applied Physics* **89** 440 (2001)

[101] S. Aghion et al. (AEḡIS collaboration), *Physical Review A* **94** 012507 (2016)

[102] R. Mayer and A. Weiss, (1988) "Comparative study of secondary-electron emission from positron and electron bombardment of Ni, Si and MgO" *Phys. Rev. B. 38, 16 (1988) 11927-11930*

[103] R.H. Howell, I.J. Rosenberg and M.J. Fluss, *Physical Review B* **35** 5303 (1987)

[104] M. Doser et al. (AEḡIS collaboration), *Classical and Quantum Gravity* **29** 184009 (2012)

[105] S. Aghion et al. (AEḡIS collaboration), *Nuclear Instruments and Methods in Physics Research B* **362** 86 (2015)

[106] T.J. Murphy, C.M. Surko, *Physical Review A* **46** 5696 (1992)

[107] D.B. Cassidy, S.H.M. Deng, H.K.M. Tanaka, and A.P. Mills, *Applied Physics Letters* **88** 194105 (2006)

[108] F. Villa, "Laser system for positronium excitation to Rydberg levels for Aegis experiment" Ph.D. Thesis, Università degli studi di Milano (2011)

[109] M. Charlton and J.W. Humberston, "Positron Physics", Cambridge University Press (2001)

[110] A.M. Alonso et al., *Physical Review A* **93** 012506 (2016)

[111] A.P. Mills Jr. et al., Nuclear Instruments and Methods in Physics Research B **192** 102 (2002)

[112] D.B. Cassidy and S.D. Hogan, *International Journal of Modern Physics: Conference Series* **30** 1460259 (2014)

[113] M.K. Oberthaler, *Nuclear Instruments and Methods in Physics Research B* **192** 129 (2002)

[114] S. Sala, *Journal of Physics B: Atomic, Molecular and Optical Physics* **48** 195002 (2015)

[115] D.A. Cooke et al., *Hyperfine Interactions* **233** 67 (2015)

[116] S. Hatamian et al., *Physical Review Letters* **58** 1833 (1987)

[117] R. Ley et al., Hyperfine Interactions **89** 327 (1994)

[118] S. Aghion et al. (AEgIS collaboration), "Producing long-lived 2^3S Ps state production via 3^3P laser excitation in magnetic and electric fields" *Awaiting publication*

[119] G. D'Agostini, *Reports on Progress in Physics* **66** 1383 (2003)

[120] S. Aghion et al. (AEgIS collaboration), "Compression of mixed antiproton and electron non-neutral plasma to high densities" *Awaiting publication*

[121] M. Hori et al., *Nuclear Instruments and Methods in Physics Research A* **496** 102 – 122 (2003)

[122] K. Hagiwara, et al. (Particle Data Group), *Physical Review D* **66** 010001 (2002)

[123] M.C. Fujiwara et al. (ATHENA collaboration), *Physical Review Letters* **92** 6 (2004)

[124] S. Aghion et al. (AEgIS collaboration), *Nuclear Instruments and Methods in Physics Research A* **732** 437 – 441 (2013)

[125] G. D'Agostini, "Bayesian reasoning in Data Analysis – A critical introduction" World scientific (2003)

[126] G. D'Agostini, "A defense of Columbo (and of the use of Bayesian inference in forensics)– A multilevel introduction to probabilistic reasoning" arXiv:1003.2086 (2010)

[127] J.D. Cockroft, E.T.S. Waltron, *Proceeding of the Royal Society of London A* **136** 619 – 630 1932

2 µm

1 µm

All items to scale

Red Light *(700nm)*

Violet Light *(400nm)*

Virus
(Enterobacteria Phage T4)

Pollen Grain
(Myosotis Alpestris)

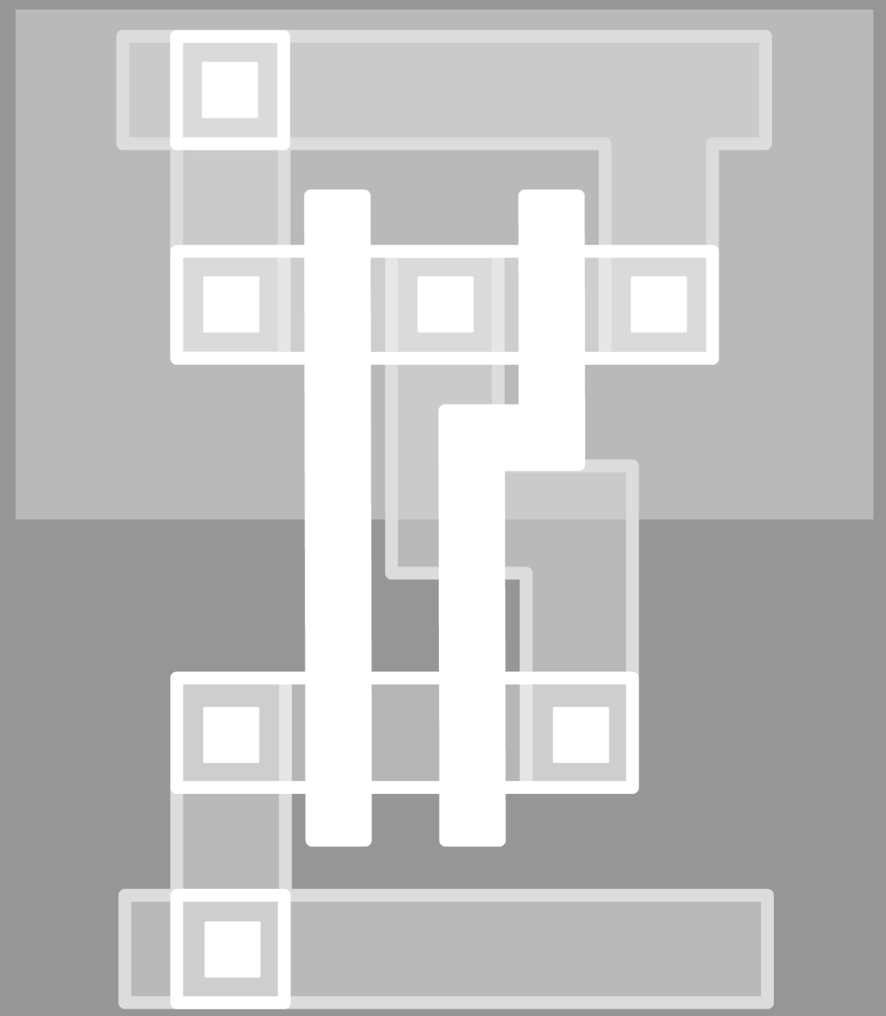
Silicon
Nanochanneled Plate
(NCP)

2 µm

1 µm

All items to scale

Bacteria
(Clostridium Botulinum)



NAND Gate
(2000)

NAND Gate
(2017)

Silicon
Nanochanneled Plate
(NCP)